

University of Groningen

De selectiviteit, efficiency en effectiviteit van het risicogerichte toezicht op grond van de WOT Winter, Heinrich; Struiksmā, Niko; Hollander, Mark

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

2015

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Winter, H., Struiksmā, N., & Hollander, M. (2015). *De selectiviteit, efficiency en effectiviteit van het risicogerichte toezicht op grond van de WOT*. Rijksuniversiteit Groningen / Pro Facto.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.



De selectiviteit, efficiency en effectiviteit van het risicogerichte toezicht op grond van de WOT

Januari 2015

Inhoud

1	Inleiding	3
1.1	Aanleiding	3
1.2	Vraagstelling	4
1.3	Aanpak	4
1.4	Dit rapport	5
2	De systematiek van het risicogerichte toezicht: kaders en normen	6
2.1	Oorsprong en hoofdlijn	6
2.2	Het theoretische model	6
2.3	Toezichtskaders	7
	2.3.1 Inleiding	7
	2.3.2 Werkwijze	8
	2.3.3 Waarderingskaders en beslisregels	9
3	Zelfevaluatie	11
3.1	Inleiding	11
3.2	Selectiviteit risicoanalyse	11
	3.2.1 Inleiding	11
	3.2.2 Primair onderwijs	11
	3.2.2 Voortgezet onderwijs	12
	3.2.3 Conclusie	13
3.3	Efficiency van het toezicht	13
	3.3.1 Inleiding	13
	3.3.2 Toezichtskosten voor de Inspectie	13
	3.3.3 Toezichtslasten voor scholen	14
	3.3.4 Conclusie	14
3.4	Effectiviteit van het toezicht	14
	3.4.1 Het aantal (zeer) zwakke scholen	14
	3.4.2 Doorlooptijd	15
	3.4.3 Onderwijsproces	16
	3.4.4 Bestuursinsteek	17
	3.4.5 Conclusie	17
4	Evaluatie van het risicogerichte toezichtsmodel	18
4.1	Inleiding	18
4.2	Het functioneren van het toezichtsmodel	18
4.3	Selectiviteit	19
4.4	Efficiency	21
4.5	Effectiviteit	22
4.6	Neveneffecten	23
4.7	Tevredenheid	24
4.8	Inspectie van het Onderwijs intern	24
4.8	Algemeen: concluderend	25
5	Conclusie	27

1 Inleiding

1.1 Aanleiding

De Wet op het onderwijstoezicht (WOT) regelt de wijze waarop in Nederland toezicht wordt gehouden op de kwaliteit van het onderwijs. Dit toezicht wordt uitgeoefend door de Inspectie van het Onderwijs (hierna: de Inspectie). Dit toezicht staat in de praktijk bekend onder de noemer 'risicogericht toezicht'. Tijdens de behandeling in de Eerste Kamer van het wetsvoorstel tot wijziging van de WOT op 24 januari 2012 heeft de minister van OCW toegezegd in 2014 een evaluatie van het risicogerichte toezicht aan het parlement te sturen. Het onderzoek richt zich op twee aspecten, te weten de effectiviteit van het toezicht en de efficiency ervan. Het ultieme einddoel van het toezicht door de Inspectie is dat de kwaliteit van het onderwijs verbetert. Als dat gebeurt, is het toezicht effectief. De efficiency van het toezicht betreft de tijd, het geld en de menskracht die het toezicht de Inspectie kost én de toezichtslasten voor de scholen. Des te beperkter de kosten voor de Inspectie en de toezichtslasten voor de scholen zijn, des te efficiënter is het toezicht.

Kern van het toezicht is dat scholen worden beoordeeld op de kwaliteit van het onderwijs en dat dat oordeel openbaar wordt gemaakt. Scholen die niet aan de normen voldoen, worden gekwalificeerd als zwak of zeer zwak. Deze scholen krijgen maximaal twee jaar de tijd de kwaliteit van het onderwijs weer op een aanvaardbaar niveau te brengen. In deze verbeterfase voert de inspectie geïntensiveerd toezicht uit.

Sinds 2007 werkt de Inspectie met het systeem van risicogericht toezicht. Dat wil zeggen dat de aard en de omvang van het toezicht worden afgestemd op de risico's op een opleiding of school. Bij de risicoanalyse worden onder meer CITO-scores, examenresultaten, in-, door-, en uitstroomgegevens, jaarstukken, signalen (zoals klachten) en nalevingsaspecten geanalyseerd. Op grond van deze analyse wordt bepaald of een school risico's loopt ten aanzien van de kwaliteit van het onderwijs. Op basis hiervan worden het niveau en de intensiteit van het toezicht bepaald. Het toezicht is er op gericht te waarborgen dat de onderwijsprestaties van scholen zich op een voldoende niveau (gaan) bevinden. Dan geldt als arrangement het basistoezicht, hetgeen inhoudt dat een school eens in de vier jaar door de inspectie wordt bezocht.

Het risicogerichte toezicht spreekt het bestuur van scholen aan. De gedachte is dat het bestuur actief, gericht en effectief optreedt waar het gaat om de kwaliteit van het onderwijs. Het schoolbestuur wordt gecontroleerd door een intern toezichthouder en treedt in dialoog met de omgeving, waaronder ouders en andere direct belanghebbenden.

De Inspectie heeft onder de noemer 'Toezicht 2020' de ambitie om te komen tot een bredere kwaliteitsverbetering van het onderwijs door ook scholen die voldoende presteren, te stimuleren zich te ontwikkelen in de richting van excellentie.

Toen in de jaren negentig van de vorige eeuw bleek dat haar oudste zoon op de openbare Montessorischool in het centrum van Amsterdam een grote leerachterstand had opgelopen, ging Karina Schaapman de strijd aan met de gemeente Amsterdam om vergoeding van noodzakelijke bijlessen. Onderzoek toonde vervolgens aan dat meer kinderen op de betreffende school leerachterstanden hadden opgelopen. Kern van het probleem was dat de school, de gemeente en de Inspectie geen idee hadden van de gerealiseerde onderwijskwaliteit en de daarop toe te passen normen. Het geschil tussen de gemeente Amsterdam en Karina Schaapman werd uiteindelijk in hoger beroep in een vonnis van 26 mei 1999 van de rechtbank Amsterdam in het voordeel van Schaapman beslecht. Dat leidde ertoe dat meer ouders bijleskosten vergoed kregen en dat meer kinderen bijgespijkerd werden. In 2000 publiceerde Karina Schaapman een boek over de affaire: 'Schoolstrijd. Ouders op de bres voor beter onderwijs.' De casus leidde uiteindelijk tot ingrijpende veranderingen in het onderwijstoezicht en de regels over governance binnen het onderwijs.

Tot 2007 bestond het onderwijstoezicht uit periodieke kwaliteitsonderzoeken en publieke rapportage over de bevindingen. Het vernieuwde toezicht streeft ernaar selectiever te zijn. Beoogd wordt te waarborgen dat het onderwijs aan de standaarden voldoet, terwijl tegelijkertijd de belasting voor het onderwijsveld en de Inspectie zelf zo beperkt mogelijk blijft. Voor de te volgen methodiek zijn binnen het primair onderwijs (hierna: PO) en het voortgezet onderwijs (hierna: VO) toezichtskaders ontwikkeld. Tot 2013 was sprake van een geïntegreerd toezichtskader PO/VO. Vanaf 2013 geldt voor het VO een herzien kader. Het toezicht op het PO werkt nog volgens het meest recente (voordien geïntegreerde) toezichtskader.

1.2 Vraagstelling

De vraagstelling die in het onderzoek centraal staat luidt:

Functioneert het model van risicogericht toezicht door de Inspectie van het Onderwijs naar behoren, gelet op de selectiviteit, efficiency en effectiviteit daarvan en wat kan over de verdere ontwikkeling van het model worden gezegd?

Deze hoofdvraag is uiteengelegd in een aantal deelvragen die de thema's selectiviteit (slaagt de risicoanalyse er in (zeer) zwakke scholen te identificeren?), efficiency (wordt de risicogerichte aanpak toegepast met zo beperkt mogelijke kosten voor de Inspectie en lasten voor de scholen?) en effectiviteit (leidt de risicogerichte aanpak tot verbetering van kwaliteit?) betreffen.

De hoofdvraag heeft betrekking op het primair onderwijs en het voortgezet onderwijs. Dat betekent dat het speciaal onderwijs buiten dit onderzoek is gebleven. De reden daarvoor is dat het risicogerichte toezicht pas in 2012 in het speciaal onderwijs is geïntroduceerd. Omdat daarover nog onvoldoende cijfers beschikbaar zijn is een evaluatie van de werking van het toezichtmodel in het speciaal onderwijs op dit moment nog niet mogelijk.

1.3 Aanpak

In dit rapport wordt de werking van het systeem van onderwijstoezicht geëvalueerd. De eerste fase in deze evaluatie betrof het opstellen door de Inspectie van een zelfevaluatie. Deze zelfevaluatie, die samen met dit rapport aan de Eerste Kamer wordt gestuurd, bevatte bouwstenen voor de uiteindelijke evaluatie die is uitgevoerd door een extern onderzoeksteam onder leiding van prof.dr. Heinrich Winter. Hij is sinds april 2014 hoogleraar Bestuurskunde aan de Rijksuniversiteit Groningen en daarvoor (vanaf mei 2010) bijzonder hoogleraar Toezicht namens Vide, de beroepsvereniging van toezichthouders, inspecteurs, handhavers en evaluatoren. Daarnaast is hij directeur van het aan de Rijksuniversiteit Groningen gelieerde onderzoeksbureau Pro Facto. Van het onderzoeksteam maakten verder mr. Niko Struiksmā (directeur Pro Facto) en mr. Mark Hollander (onderzoeker Pro Facto) deel uit.

De onderzoekers hebben de volgende activiteiten verricht:

- analyse zelfevaluatie
- uitgebreide document- en literatuurstudie (zie bijlage 1)
- interviews met:
 - medewerkers Onderzoek en Informatievoorziening van de Inspectie
 - twee inspecteurs PO
 - twee inspecteurs VO
 - twee schoolbestuurders VO waarvan één van een zwakke school
 - één schoolbestuurder PO van een zwakke school
 - vier directeuren VO waarvan drie van (zeer) zwakke scholen
 - vier directeuren PO waarvan drie van (zeer) zwakke scholen

Dit rapport vormt de weerslag van de genoemde activiteiten. Daarbij is gebruik gemaakt van statistische informatie en analyses uit de zelfevaluatie. Deze zijn opgenomen in hoofdstuk 3 van dit rapport. Onze eigen analyse en beoordeling daarvan vinden plaats in de daaropvolgende hoofdstukken, waarbij bij het beoordelen van het functioneren van het risicogerichte toezicht ook gebruik is gemaakt van andere informatiebronnen dan de zelfevaluatie.

1.4 Dit rapport

In hoofdstuk 2 wordt het model van het risicogerichte toezicht beschreven. Daarbij komen de verschillende toezichtskaders voor het primair en het voortgezet onderwijs aan de orde. In hoofdstuk 3 staat de zelfevaluatie van de Inspectie centraal. In hoofdstuk 4 komen we tot een evaluatie van het systeem van onderwijstoezicht. De rapportage sluit in hoofdstuk 5 af met een beschouwing waarin ook aandacht wordt besteed aan de ambities van de Inspectie met ‘Toezicht 2020’.

2 De systematiek van het risicogerichte toezicht: kaders en normen

2.1 Oorsprong en hoofdlijn

In 2005 verscheen de Tweede Kaderstellende Visie op Toezicht.¹ Dit document geldt nog steeds als basis voor het toezicht zoals uitgeoefend door rijksinspecties en autoriteiten. In de Kaderstellende Visie zijn zes principes van goed toezicht geformuleerd. Eén van die principes is dat toezicht selectief moet zijn. Volgens dat uitgangspunt is het een taak van de overheid na te gaan in hoeverre burgers, bedrijven en instellingen zelf verantwoordelijk kunnen zijn bij het handhaven van regels. Toezichthouders moeten daarbij een afweging maken op basis van de risico's en de kosten en baten van toezicht. Toezicht moet zoveel mogelijk aansluiten bij andere vormen van controle en verantwoording. Waar de toezichthouder meer ruimte laat en vertrouwen geeft, kan het toezicht een stapje terug doen.

Het beginsel van selectief toezicht is vertaald in de Visie op Toezicht van het ministerie van OCW.² Voor het onderwijsveld betekent het dat de onderwijsinspectie terugvalt op de mechanismen binnen scholen die zorgen voor het waarborgen van de onderwijskwaliteit. Het is in de eerste plaats het schoolbestuur dat door de Inspectie wordt aangesproken op zijn verantwoordelijkheid voor onderwijskwaliteit. De Inspectie noemt dit 'de bestuursinsteek'. Selectief toezicht betekent in de tweede plaats de afstemming van aard en intensiteit van het toezicht op de risico's die volgens de toezichtsmethodiek van de Inspectie op scholen aanwezig zijn. Deze werkwijze wordt doorgaans aangeduid als het 'risicogerichte toezichtmodel'. De introductie van dit model zou moeten leiden tot een reductie van de toezichtslast voor de scholen met 25%.³

Bij het risicogerichte toezichtmodel bepaalt een risicoanalyse de aard en de intensiteit van het toezicht. Bij deze risicoanalyse staan 'opbrengsten' centraal. Daarmee wordt bedoeld op CITO-scores, eindexamencijfers en rendementscijfers. Wanneer de risicoanalyse duidt op te lage opbrengsten wordt door de Inspectie een schoolbezoek afgelegd, waarbij onderzoek wordt gedaan naar het leerproces, leerlingenzorg en begeleiding en kwaliteitszorg. Scholen waar geen risico's zijn vastgesteld krijgen geen toezichtsonderzoek, met dien verstande dat elke school tenminste eens per vier jaar wordt bezocht. Dit vierjaarlijkse onderzoek controleert de werking van de systematiek van risicoanalyse.

Voor het primair onderwijs en het voortgezet onderwijs is dit toezichtmodel uitgewerkt in een toezichtskader en een waarderingskader. Die worden hierna onder paragraaf 2.3 beschreven. Eerst gaan we in op het theoretische model.

2.2 Het theoretische model

De beleidstheorie van het risicogerichte toezicht en de Wet op het Onderwijstoezicht onderscheidt een aantal verschillende actoren: de overheid/de toezichthouder, het autonome bestuur van de school en diverse belanghebbenden, zoals ouders en leerlingen. De beleidstheorie van het onderwijstoezicht gaat uit van een aantal aannames over het functioneren van het toezicht. Wensen en klachten van belanghebbenden vormen de (mogelijke) input voor het selectieve toezicht, zeker als de school niet reageert wanneer die klachten en wensen al eerder zijn ingebracht. Het bevoegd gezag, doorgaans het schoolbestuur (de bestuurder of de

¹ TK 2005/06, 27 831, nr. 15. 'Minder last, meer effect'.

² Toezicht in vertrouwen, vertrouwen in toezicht. Visie op toezicht 2007-2011, OCW september 2006.

³ Deze doelstelling is terug te vinden in 'Evaluatie risicogericht toezicht', Bijlage bij TK 2009/10, 32 123 VIII, nr.127, p. 7, waar wordt verwezen naar de Kaderstellende visie op toezicht van 2005, de visie op toezicht van OCW uit 2006, het coalitieakkoord en de evaluatie van de WOT uit 2007.

raad van bestuur) staat centraal in het toezichtsmodel. Het bestuur is als enige bevoegd tot het nemen van maatregelen gericht op de kwaliteit van het onderwijs en is daarvoor dus ook verantwoordelijk. Het bestuur verantwoordt zich extern voor die kwaliteit. Dat betekent dat aan de belanghebbenden, maar ook aan de toezichthouder informatie over de onderwijskwaliteit wordt verstrekt. Dat het schoolbestuur zich extern moet verantwoorden leidt, zo wordt verondersteld, tot versterking van het interne toezicht en tot de ontwikkeling van een model voor kwaliteitszorg. De externe verantwoordingsinformatie wordt volgens het model door de Inspectie gevalideerd met behulp van onafhankelijke bronnen. Op grond daarvan wordt vervolgens de risicoanalyse uitgevoerd. Voor het beoordelen van scholen hanteert de Inspectie toezichtskaders die alle relevante aspecten van het te beoordelen onderwijs beogen te omvatten en die in overleg met het onderwijsveld tot stand zijn gekomen.

2.3 Toezichtskaders

2.3.1 Inleiding

De Inspectie stelt toezichtskaders vast om de eenheid van handelen van de inspecteurs op de werkvloer te bevorderen. Het voordeel van een geëxpliciteerd kader is ook dat de werkwijze van de Inspectie kenbaar is voor betrokkenen en belanghebbenden. In 2009 zijn de toezichtskaders PO en VO geactualiseerd. In 2009 is gekozen voor een geïntegreerd toezichtskader PO/VO. Het toezichtskader bevat twee elementen: de werkwijze die de Inspectie hanteert (het hoe) en de kwaliteitsaspecten en indicatoren (het waarderingskader of kernkader) waarop scholen worden beoordeeld (het wat). Het kernkader PO omvat op dit moment 54 indicatoren en het kernkader VO 46 indicatoren waar bij een schoolbezoek op gelet wordt.

Het toezichtskader PO/VO van 2009 is in 2011 aangepast (onder meer naar aanleiding van de in de Wet ‘Goed onderwijs, goed bestuur’ opgenomen bevoegdheid tot het geheel of gedeeltelijk beëindigen van de bekostiging van een school door de minister en de komende aanpassing van de standaard-verbetertermijn voor zeer zwakke scholen van twee naar 1 jaar). In juli 2012 zijn enkele technische wijzigingen doorgevoerd in het toezichtskader PO/VO. Dat kader is vanaf het schooljaar 2012/2013 gehanteerd.

Op 20 september 2012 is met name op het onderdeel waarderingskader een grondige aanpassing van het toezichtskader 2011 vastgesteld voor het VO.⁴ Vanaf 1 januari 2013 geldt het toezichtskader PO/VO 2012 voor het PO en vanaf datzelfde moment het toezichtskader VO 2013 voor het VO, behalve voor VO-scholen die voor 1 januari 2013 in een aangepast toezichtstraject zaten. Deze scholen worden nog volgens het geïntegreerde toezichtskader benaderd. Er is dus momenteel sprake van twee toezichtskaders, één voor PO en één voor VO.

De werkwijze van de Inspectie kent vier uitgangspunten:

- het toezicht kent een bestuursgerichte aanpak;
- het toezicht is risicogericht en op maat;
- het toezicht is er mede op gericht te voorkomen dat scholen zwak of zeer zwak worden;
- het toezicht is mede gericht op (programmatisch) handhaven.

Met de invoering van de WOT is een kernthema van het toezicht het aanspreken van het bevoegde gezag op zijn verantwoordelijkheid voor de onderwijskwaliteit. Dat betreft in eerste instantie de behaalde onderwijsresultaten en, pas bij onvoldoende resultaten, de verrichte inspanningen. Daarbij worden scholen niet meer belast dan nodig is (art. 4, lid 2 WOT).

⁴ Het waarderingskader wordt regelmatig op onderdelen aangepast. Zo werd op 5 september 2013 een beleidsregel van de Inspecteur-generaal van het Onderwijs, nr. 4160445, gepubliceerd, houdende het instellen van een beslisregel bij het Toezichtkader VO 2013 en het van toepassing verklaren van het Toezichtkader VO 2013 op scholen als bedoeld in artikel 56 Wet op het voortgezet onderwijs. Op 19 december 2013 is het waarderingskader opnieuw op onderdelen aangepast.

2.3.2 Werkwijze

De werkwijze van de Inspectie is erop gericht om op basis van de opbrengsten, jaarstukken en signalen een analyse te maken voor alle onderwijsinstellingen, ter bepaling van de mate van intensiteit van het toezicht. De mate van intensiteit varieert met de waargenomen risico's: toezicht op maat dus. Wanneer sprake is van risico's worden aanvullende gegevens opgevraagd en geanalyseerd. Indien nodig wordt onderzoek naar de onderwijskwaliteit uitgevoerd door middel van een schoolbezoek. Worden dan tekortkomingen vastgesteld op basis van zogenaamde beslisregels (zie hierna), dan wordt een geïntensiveerde vorm van toezicht in een toezichtsarrangement vastgelegd. Dat kan zijn het aangepaste arrangement 'zwak' of het arrangement 'zeer zwak'. Een (zeer) zwakke school krijgt maximaal twee jaar de tijd zich te verbeteren en in het basisarrangement te komen.

De Inspectie staan verschillende interventies ten dienste om te voorkomen dat een school zwak of zeer zwak wordt. Zijn de opbrengsten in een jaar onder de norm, dan wordt het schoolbestuur daarvan op de hoogte gesteld door middel van een zogenaamde attendering. Soms wordt een gesprek met het schoolbestuur gevoerd. Zijn de opbrengsten in het tweede daaropvolgende jaar nog steeds onvoldoende, dan geeft de Inspectie een waarschuwing af aan het bestuur en wordt om een plan van aanpak verzocht. Er is dan altijd een gesprek met het schoolbestuur en soms een schoolbezoek. Tot slot wordt in het derde jaar, wanneer de opbrengsten voor het derde opeenvolgende jaar onvoldoende zijn en nadat een schoolbezoek is afgelegd, een aangepast toezichtsarrangement vastgesteld. De school wordt dan zwak of zeer zwak. Blijven ook dan nog verbeteringen uit, dan zijn vervolgens andere maatregelen mogelijk, zoals het intrekken van (een deel van) de bekostiging of het opheffen van een school.

Er zijn dus drie fasen in het toezichtsproces te onderscheiden:

- het uitvoeren van een risicoanalyse;
- het uitvoeren van een onderzoek naar mogelijke tekortkomingen;
- het uitvoeren van een interventietraject.

Voldoen de opbrengsten aan de norm, dan ontvangt de school een basisarrangement. Zijn de opbrengsten onder de norm, dan wordt vervolgens – wanneer die onvoldoende opbrengsten aanhouden – overleg gevoerd met het schoolbestuur. Daarbij spelen verantwoordingsdocumenten een belangrijke rol. Die documenten laten zien dat van intern toezicht en horizontale verantwoording aan de omgeving sprake is. Vervolgens wordt in de meeste gevallen een kwaliteitsonderzoek, een schoolbezoek, uitgevoerd aan de hand van indicatoren uit het zogenaamde kernkader (zie hierna). Het schoolbezoek wordt ruim van tevoren (minimaal vier weken) schriftelijk aangekondigd. De inspecteurs voeren ter plaatse documentstudie uit, observeren lessen en voeren gesprekken met leerlingen, ouders, leraren en leidinggevenden. De bevindingen worden aan het eind van het bezoek teruggekoppeld aan het bestuur en de schoolleiding. Binnen zeven weken na afloop van het bezoek wordt een rapport van bevindingen opgesteld met daarin een toezichtsarrangement. Na hoor- en wederhoor wordt het rapport door de inspectie vastgesteld.

Voor het vaststellen van een toezichtsarrangement waarmee een school zwak of zeer zwak wordt verklaard, worden beslisregels toegepast. Daarbij spelen de eindopbrengsten een belangrijke rol. De Inspectie kan van de beslisregels afwijken als de opbrengsten voldoende zijn, maar niettemin grote risico's worden gesignaleerd. Ook wanneer de opbrengsten onvoldoende zijn, kan de Inspectie besluiten toch het basisarrangement toe te kennen als de overige indicatoren voldoende zijn en er ook een verklaring is voor het niveau van de opbrengsten. Dat gebeurt een zeer beperkt aantal keren per jaar.

Voordat aan een school of opleiding het oordeel zwak wordt toegekend, hebben binnen de Inspectie een referent en de teamleider meegekeken. Het oordeel is dus niet slechts gebaseerd op de analyse van de opbrengsten en het oordeel van de inspecteurs op basis van het schoolbezoek. Wordt het oordeel zeer zwak toegekend, dan is er behalve een rol voor de teamleider en een referent ook overleg met een landelijke toetsingscommissie van de Inspectie.

Behalve bezoeken aan scholen die op basis van een analyse van de opbrengsten (zeer) zwak dreigen te worden, worden ook reguliere schoolbezoeken afgelegd. Elke school wordt in beginsel eens in de vier jaar bezocht. Tijdens dit bezoek wordt in principe geen gesprek met het schoolbestuur gevoerd.

2.3.3 Waarderingskaders en beslisregels

De grondslagen voor het waarderingskader zijn te vinden in de sectorwetten, de Wet op het primair onderwijs en de Wet op het voorgezet onderwijs en de WOT 2012.

De waarderingskaders van het toezichtskader VO 2013 en van het toezichtskader PO 2012 zijn gebaseerd op artikel 11, lid 2 en 3, WOT. Van belang zijn (lid 2):

- a. leerresultaten;
- b. voortgang in de ontwikkeling van leerlingen;
- c. het gevoerde personeelsbeleid, voor zover daar op grond van artikel 6a, eerste en tweede lid, aanleiding toe bestaat.

Als er een vermoeden bestaat dat het onderwijs tekortschiet, stelt de inspectie nader onderzoek in, waarbij tevens de oorzaken van het tekortschieten worden onderzocht. Bij dit nader onderzoek worden volgens lid 3 van artikel 11 WOT een of meer van de volgende aspecten van kwaliteit betrokken, te weten:

- a. het leerstofaanbod
- b. de leertijd
- c. het pedagogisch klimaat
- d. het schoolklimaat
- e. de effectiviteit van het didactisch handelen van het onderwijspersoneel
- f. de leerlingenzorg
- g. de inhoud, het niveau en de uitvoering van de toetsen, tests, opdrachten of examens
- h. voor wat betreft de kwaliteit van het onderwijspersoneel:
 - 1°. de eisen van bevoegdheid en bekwaamheid,
 - 2°. het personeelsbeleid, voor zover het betreft de duurzame borging van de kwaliteit van het onderwijspersoneel”.

Daarnaast vermeldt de WOT dat de kwaliteitszorg van de school bij het toezicht betrokken wordt (artikel 12, eerste lid, WOT).

Waarderingskader PO

Het waarderingskader van de Inspectie bestaat uit 54 indicatoren voor het PO. Het waarderingskader PO is onderverdeeld in vijf domeinen:

- opbrengsten,
- onderwijsleerproces,
- (leerlingen)zorg en begeleiding,
- kwaliteitszorg en
- wet- en regelgeving.

Binnen het domein opbrengsten is sprake van twee normindicatoren: de beoordeling van de eindresultaten (in de praktijk vrijwel uitsluitend: de CITO-score) en de beoordeling van de tussenresultaten. De eerste indicator betreft de resultaten van leerlingen aan het eind van de basisschool; die moeten ten minste op het niveau liggen dat op grond van kenmerken van de leerlingenpopulatie mag worden verwacht. De tweede indicator betreft de resultaten van de leerlingen voor Nederlandse taal en rekenen en wiskunde tijdens de schoolperiode; die moeten ten minste op het niveau liggen dat op grond van de kenmerken van de leerlingenpopulatie mag worden verwacht.

Voor de domeinen onderwijsleerproces en leerlingenzorg worden acht normindicatoren gehanteerd. Wanneer een school niet beschikt over eindresultaten van leerlingen en niet over

tussentijdse resultaten worden naast deze acht indicatoren de zes indicatoren van het domein kwaliteitszorg bij de normering betrokken.

Daarbij gelden de volgende beslisregels: een school is zeer zwak als er op basis van de indicator voor de eindresultaten onvoldoende opbrengsten zijn en twee of meer normindicatoren van het onderwijsleerproces en de leerlingenzorg onvoldoende zijn. Zijn de eindopbrengsten niet te beoordelen, dan wordt een school als zeer zwak gekwalificeerd als meer dan de helft van de (beoordeelde) indicatoren voor kwaliteitszorg onvoldoende is en twee of meer normindicatoren van het onderwijsleerproces en de leerlingenzorg onvoldoende zijn. Daarnaast geldt dat de Inspectie altijd de kwalificatie (zeer) zwak kan toekennen als er grote risico's worden gezien rond de kwaliteit van de school, ook als de eindopbrengsten voldoende zijn, zoals bij een combinatie van factoren als tussentijdse leerprestaties die niet aan de norm voldoen en/of drie of meer normindicatoren voor het onderwijsleerproces die onvoldoende zijn.⁵

Waarderingskader VO

Het waarderingskader VO bestond aanvankelijk uit dezelfde vijf domeinen als bij PO, maar sinds de actualisatie van 2013 is sprake van drie domeinen:

- opbrengsten,
- onderwijsproces, en
- kwaliteitszorg.

Er wordt in het waarderingskader VO gesproken over drie soorten indicatoren: het kernkader, indicatoren die de Inspectie in een onderzoek altijd wil onderzoeken, het aanvullend kader, als er op basis van het bureauonderzoek reden is de betreffende indicatoren te beoordelen en het verdiepend kader wanneer de Inspectie een beter inzicht wil krijgen in de oorzaken van achterblijvende kwaliteit. Het kernkader VO bestaat uit 21 indicatoren.

Voor het beoordelen van de toepasselijkheid van een toezichtsarrangement worden in ieder geval de vier indicatoren uit het domein opbrengsten beoordeeld. Daarbij gaat het om:

- de resultaten van het centraal eindexamen,
- de verschillen tussen schoolexamens en het centraal eindexamen
- het rendement (doorstroming) in de onderbouw
- het rendement (doorstroming) in de bovenbouw

Zijn twee of meer indicatoren onvoldoende, dan zijn de opbrengsten onvoldoende. Voor de indicatoren uit het domein onderwijsproces geldt de beslisregel dat de kwaliteit onvoldoende is als drie of meer van de acht indicatoren uit dit domein onvoldoende zijn. De kwaliteitszorg is volgens de beslisregel onvoldoende als twee of meer van de vier indicatoren uit het kernkader kwaliteitszorg onvoldoende zijn. Indien niet of nauwelijks sprake is van tekortkomingen wordt het basisarrangement toegekend.

Voor zowel PO als VO geldt dat wanneer drie jaren achtereen de opbrengsten als onvoldoende zijn beoordeeld, de school als (zeer) zwak wordt gekwalificeerd en in een aangepast toezichtsarrangement terecht komt.

⁵ De beslisregels zijn te vinden in de 'Beslisschema's aangepast arrangement PO'.

3 Zelfevaluatie

3.1 Inleiding

Ten behoeve van de evaluatie van het risicogerichte toezicht heeft de Inspectie een zelfevaluatie opgesteld. Die betreft onder meer de selectiviteit, de efficiency en de effectiviteit van het risicogerichte toezicht. In dit hoofdstuk wordt die zelfevaluatie besproken. In enkele gevallen wordt daarbij aanvullende informatie betrokken die naar aanleiding van de zelfevaluatie door de onderzoekers is opgevraagd. In het volgende hoofdstuk wordt het risicogerichte toezicht geëvalueerd. Op die plaats komen elementen van de zelfevaluatie opnieuw aan de orde. Achtereenvolgens worden hier eerst besproken de selectiviteit (par. 3.2), de efficiency (par. 3.3) en de effectiviteit van de risicoanalyse (par. 3.4).

3.2 Selectiviteit risicoanalyse

3.2.1 Inleiding

Om de betrouwbaarheid van de risicoanalyse te toetsen, heeft de Inspectie de risico-inschatting, gebaseerd op analyses van de opbrengsten, vergeleken met de beoordeling van de onderwijskwaliteit door de inspecteurs op basis van schoolbezoeken. Het ingeschatte risiconiveau is vergeleken met het arrangement zoals vastgesteld na afloop van een schoolbezoek.

3.2.2 Primair onderwijs

In tabel 3.1 zijn de uitkomsten weergegeven van de risicoanalyse op basis van opbrengsten en het uiteindelijke arrangement dat de inspectie zou moeten toekennen in onderzoeken over 2007 en 2008 bij in totaal 640 scholen voor primair onderwijs. Een voorbeeld ter illustratie: bij 281 scholen zijn geen risico's gesignaleerd en is het uiteindelijk toegekende arrangement 'basis'. Het gaat hier om een aselechte steekproef van scholen die allemaal door een inspecteur onderzocht en beoordeeld zijn.

TABEL 3.1: RISICOANALYSE VERSUS FEITELIJK ARRANGEMENT 2007/2008 (N=640)

	Toegekend arrangement		
	Basis	Zwak	Zeer zwak
Geen risico (N=286)	98,3% (N=281)	1,4% (N=5)	0,3% (N=1)
Risico (N=303)	93,4% (N=283)	5,6% (N=17)	1,0% (N=3)
Ernstig risico (N=51)	51,0% (N=26)	41,2% (N=21)	7,8% (N=4)

'Geen risico' betekent in dit verband dat de scholen positief scoren op de opbrengsten, bij 'risico' zijn er onvoldoende gegevens om dat vast te stellen. Bij 'ernstig risico' zijn de opbrengsten onvoldoende.

De tabel laat volgens de zelfevaluatie zien dat de risicoanalyse "zeer voorzichtig was afgesteld". Daarmee wordt bedoeld dat in relatief veel gevallen waar op grond van de risicoanalyse sprake zou (kunnen) zijn van risico, dit uiteindelijk volgens het toegekende arrangement niet het geval bleek te zijn. Hiervoor is de rij 'ernstig risico' bepalend. Van de 51 scholen die het betreft blijkt dat in ongeveer de helft van de gevallen waarin op basis van de analyse van opbrengsten sprake zou zijn van ernstig risico, de scholen toch in het basis toezichtsarrangement terecht zijn gekomen. Het gaat dus om 'false negatives'; deze scholen zijn in de risicoanalyse ten onrechte als risicovol bestempeld.

Met het voorgaande is de vraag naar ‘false positives’ nauw verbonden: in welke mate is bij scholen volgens de risicoanalyse geen sprake van risico, terwijl naar aanleiding van het reguliere schoolbezoek dat eens in de vier jaar plaats vindt, de school uiteindelijk toch als (zeer) zwak wordt gekwalificeerd. De zelfevaluatie doet daarover geen uitspraken.

In tabel 3.2 zijn de risico-analyse en het feitelijke arrangement aan elkaar gekoppeld voor het jaar 2013. Daarbij zijn, zo blijkt, andere categorieën voor de risico-inschatting gehanteerd dan in 2007/2008.

TABEL 3.2: RISICOANALYSE VERSUS FEITELIJK ARRANGEMENT 2013 (N=6991)

	Toegekend arrangement			
	Basis	Zwak	Zeer zwak	Geen arrangement
Foute verantwoording (N=8)	62,5% (N=5)	12,5% (N=1)	0,0% (N=0)	25% (N=2)
Onbekend risico (N=679)	90,9% (N=617)	4,4% (N=30)	1,3% (N=9)	3,4% (N=23)
Geen risico (N=4585)	99,1% (N=4546)	0,6% (N=28)	0,0% (N=0)	0,2% (N=11)
Enig Risico (N=1198)	99,0% (N=1186)	0,8% (N=10)	0,0% (N=0)	0,2% (N=2)
Risico (N=339)	91,4% (N=310)	7,1% (N=24)	0,6% (N=2)	0,9% (N=3)
Ernstig risico (N=182)	48,9% (N=89)	44,5% (N=81)	6,6% (N=12)	0,0% (N=0)

De vergelijking tussen tabel 3.1 en 3.2 laat zien dat de kans dat een school met ernstig risico toch in het basisarrangement terecht komt is afgenomen van iets meer dan 50% tot iets daaronder. Afgezet als gedeelte van de totale populatie is de daling van het aantal ‘false negatives’ duidelijk groter: van 4% (26 van 640) naar ruim 1% (89 van 6991).

Het aantal ‘false positives’ in tabel 3.2 is door ons berekend door de scholen in de categorieën ‘geen risico’ en ‘enig risico’ die toch als ‘zwak’ en ‘zeer zwak’ zijn gekwalificeerd te vergelijken met het totaal aantal ‘(zeer) zwakke’ scholen. Het gaat om 38 van 157: 4%. Ook hier is dus sprake van een afname ten opzichte van 2007/2008 toen dat aandeel nog 12% was.

3.2.2 Voortgezet onderwijs

In de tabellen 3.3 en 3.4 worden de feitelijke toezichtsarrangementen in het VO van 2011 en 2013 vergeleken met de risicoanalyses die zijn gemaakt in 2010 en 2012. Het gaat daarbij dus, evenals hierboven bij het primair onderwijs, om de risicodetectie op basis van een opbrengstenanalyse.

TABEL 3.3: RISICOANALYSE 2010 VERSUS FEITELIJK ARRANGEMENT 2011 (N=2527)

	Toegekend arrangement		
	Basis	Zwak	Zeer zwak
Geen risico (N=1501)	99,7% (N=1496)	0,3% (N=5)	0,0% (N=0)
Risico (N=730)	94,8% (N=692)	4,1% (N=30)	1,1% (N=8)
Ernstig risico (N=296)	11,8% (N=35)	80,7% (N=239)	7,4% (N=22)

TABEL 3.4: RISICOANALYSE 2012 VERSUS FEITELIJK ARRANGEMENT 2013 (N=2442)

	Toegekend arrangement		
	Basis	Zwak	Zeer zwak
Geen risico (N=1504)	99,7% (N=1499)	0,3% (N=5)	0,0% (N=0)
Risico (N=653)	95,1% (N=621)	4,7% (N=31)	0,2% (N=1)
Ernstig risico (N=285)	28,1% (N=80)	66,3% (N=189)	5,6% (N=16)

Opvallend is dat in vergelijking met 2010/2011, de risicoanalyse 2012/2013 een hoger percentage ‘false negatives’ laat zien. De risicoanalyse is dus strenger geworden in die zin dat in

de opbrengstenanalyse gesignaleerde risico's in de latere analyse vaker blijken mee te vallen, zodat toch een basisarrangement wordt toegekend. De Inspectie stelt hierover zelf dat die toename het gevolg kan zijn van het na 2011 gebruiken van de septemberlevering van de examengegevens. Scholen die volgens de risicoanalyse, gebaseerd op de examens van jaar T, zwak scoren, maar waarvan uit de septemberlevering van de examendata van jaar T+1 blijkt dat de onvoldoende opbrengsten van het centraal eindexamen en/of het verschil tussen schoolexamens en het centraal eindexamen mogelijk zijn opgeheven, krijgen vooruitlopend op de volgende risicoanalyse het basisarrangement toegekend.

3.2.3 Conclusie

Concluderend wordt in de zelfevaluatie gesteld dat de toetsing van de indicatoren aan de waargenomen werkelijkheid laat zien dat de keuze voor de risico-indicatoren goed lijkt. Tegelijkertijd wordt in de zelfevaluatie gesteld dat veel indicatoren vooral indicaties zijn van onderpresteren. Daarmee wordt bedoeld dat onderpresteren wordt waargenomen nadat eerst onvoldoende opbrengsten zijn vastgesteld. De zelfevaluatie koppelt dat aan de observatie dat in de gevolgde systematiek vroegtijdige signalering nog niet altijd aan de orde is.

3.3 Efficiency van het toezicht

3.3.1 Inleiding

De invoering van het risicogerichte toezicht heeft op verschillende manieren gevolgen gehad. In de oorspronkelijke inrichting van het toezicht volgens de WOT (dus voor het risicogerichte toezicht) was sprake van een bezoek per jaar, gecombineerd met ten minste één uitgebreid onderzoek in de vier jaar. In de nieuwe inrichting is het jaarlijkse bezoek vervangen door de risicoanalyse. Als scholen dan risico's of tekortkomingen vertonen, krijgen zij vervolgens te maken met een veel zwaardere belasting dan in het verleden. Scholen waar niets aan de hand is, krijgen eenmaal in de vier jaar een lichte vorm van onderzoek. In totaliteit is er sprake van een reductie van de lasten voor zowel de scholen als de Inspectie. In de volgende subparagrafen wordt nader ingegaan op de toezichtskosten voor de Inspectie en de toezichtslasten voor de scholen.

3.3.2 Toezichtskosten voor de Inspectie

In de onderstaande tabel is de ontwikkeling te zien van het aantal fte's dat werkzaam is bij de Inspectie, per jaar, uitgesplitst per sector. De formatie is geleidelijk teruggelopen als resultaat van de taakstellingen met betrekking tot de uitgaven van de rijksoverheid.

TABEL 3.5: FORMATIE PER SECTOR VOOR TOEZICHT

	PO	Vo
2010	133	46,8
2011	131	46,8
2012	131,1	46,1
2013	131,1	44,6
2014	126,8	44,6

Uit de beschikbare formatie concludeert de Inspectie dat er geen sprake kan zijn van toename van de toezichtlast voor de scholen omdat voor het toezicht minder inspecteurs beschikbaar zijn en het aantal scholen in belangrijke mate stabiel is gebleven.

Of de Inspectie de beschikbare capaciteit effectief heeft ingezet, komt aan de orde in de par. 3.4.

3.3.3 Toezichtslasten voor scholen

Een ex ante evaluatie uit 2009 door de Inspectie die is opgenomen in de zelfevaluatie, geeft aan dat er voor zowel PO- als VO-scholen sprake zou zijn van een vermindering van de toezichtslasten met 42% voor het voortgezet onderwijs en 48% voor het primair onderwijs. De toezichtslasten bestaan dan uit de lasten die samenhangen met schoolbezoeken, maar ook met verzoeken om informatie vanuit de Inspectie.

Uit de ex ante evaluatie komt naar voren dat scholen in het primair onderwijs onder de WOT 2001 jaarlijks gemiddeld 24,6 uren besteedden aan inspectietoezicht. Het risicogerichte toezicht levert alle basisscholen (goede en zwakke) volgens de zelfevaluatie een toezichtlast van gemiddeld 12,7 uur per jaar op. Voor zwakke of zeer zwakke scholen komt daar vervolgens 26,7 uur extra toezichtlast bij, zodat daar de totale last 39,4 uur per jaar bedraagt. De gemiddelde bemoeienis met goede scholen is dus bijna gehalveerd, terwijl die voor de zwakke scholen met bijna 40 procent is toegenomen. Uit de zelfevaluatie wordt niet duidelijk voor welk jaar deze cijfers precies gelden en hoe deze zich ontwikkeld hebben. Over de feitelijke ureninzet voor inspectietoezicht in het voortgezet onderwijs zijn geen gegevens beschikbaar in de zelfevaluatie.

In de zelfevaluatie zijn de toezichtslasten ook geanalyseerd door te kijken naar de bevragsingslast. Onder bevragsingslast wordt verstaan het aantal 'bevragingen' (vragenlijsten, enquêtes, andere informatieverzoeken) per jaar. Het cijfer zegt niets over de omvang van de bevragingen, alleen over het aantal. Uit de zelfevaluatie blijkt dat de bevragsingslast voor het primair onderwijs de afgelopen jaren stabiel is, in de afgelopen jaren uiteenlopend van 0,55 tot 0,6 bevragingen per jaar. Voor het voortgezet onderwijs is een licht dalende lijn te zien, van 0,4 in 2010 tot 0,32 in 2013.

In het kader van een tevredenheidsonderzoek door de Inspectie naar de toezichtslasten is de scholen gevraagd of de investering die zij zelf in het inspectieonderzoek hebben gedaan, in verhouding staat met de opbrengsten die zij ervan ervaren. In zowel het primair als in het voortgezet onderwijs was 80 procent van de scholen het er (zeer) mee eens dat dit het geval was. Voor beide sectoren is op dit punt een kleine afname in de tevredenheid zichtbaar.

3.3.4 Conclusie

Concluderend wordt in de zelfevaluatie het volgende gesteld: "Het proportionele en selectieve toezicht heeft geleid tot een belangrijke efficiëncyslag voor de inspectie en op scholen. Op de scholen is sprake van een reductie van de toezichtlast, en ook de inspectie werkt efficiënter. Tegelijkertijd zien we een verschuiving van de lasten op scholen. Gerichter toezicht leidt tot minder inspecties op goede scholen en tot een halvering van de toezichtlast op deze scholen. Toezicht op de (zeer) zwakke scholen is geïntensiveerd. Dit betekent meer last voor deze scholen."

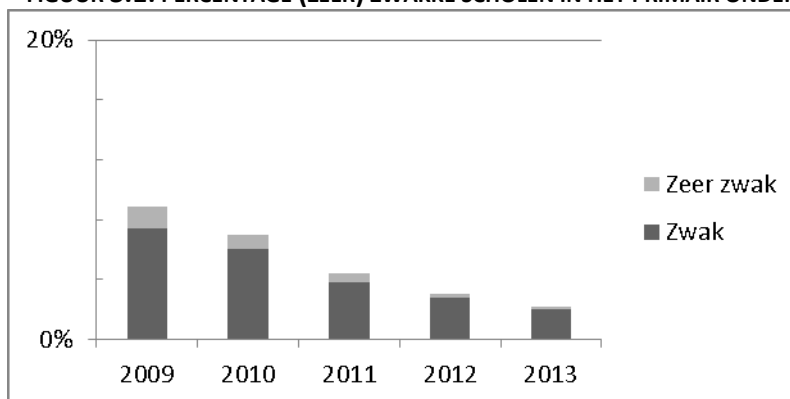
3.4 Effectiviteit van het toezicht

3.4.1 Het aantal (zeer) zwakke scholen

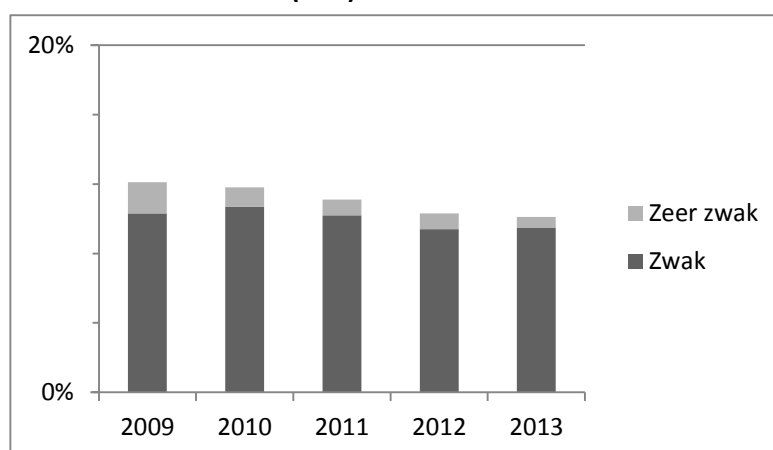
De Inspectie hanteert als uitgangspunt dat toezicht op zwakke scholen meer effect sorteert dan toezicht op goede scholen. In het kader van het risicogerichte toezicht wordt actief geïntervenieerd om de kwaliteit van (zeer) zwakke scholen te verbeteren en deze verbetering te bespoedigen. Scholen worden geacht zich in maximaal twee jaar te verbeteren van (zeer) zwak tot een niveau waarop zij weer basistoezicht krijgen.

Een van de manieren om naar de effectiviteit van het toezicht te kijken, is de ontwikkeling in het aantal (zeer) zwakke scholen te bestuderen. In de figuren 3.1 en 3.2 is de ontwikkeling van het percentage (zeer) zwakke scholen tussen 2009 en 2013 voor primair en voortgezet onderwijs weergegeven.

FIGUUR 3.1: PERCENTAGE (ZEER) ZWAKKE SCHOLEN IN HET PRIMAIR ONDERWIJS



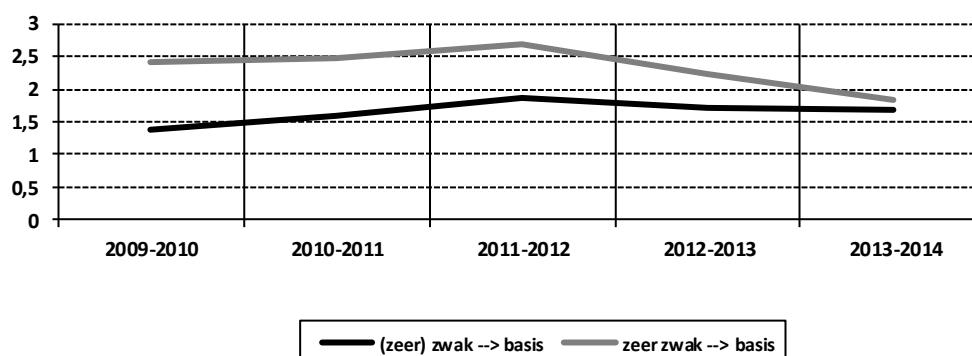
FIGUUR 3.2: PERCENTAGE (ZEER) ZWAKKE SCHOLEN IN HET VOORTGEZET ONDERWIJS



Vanaf 2009 is er bij zowel het primair als het voortgezet onderwijs sprake van een afname van het aantal (zeer) zwakke scholen, zo blijkt uit bovenstaande figuren. In het voortgezet onderwijs is de teruggang van het aantal (zeer) zwakke scholen geringer dan in het primair onderwijs. Op grond van de ontwikkeling van het aantal zwakke en zeer zwakke scholen zou, zo wordt in de zelfevaluatie gesteld, kan worden geconcludeerd dat het risicogerichte toezicht zeer effectief is geweest in het waarborgen van de basiskwaliteit.

3.4.2 Doorlooptijd

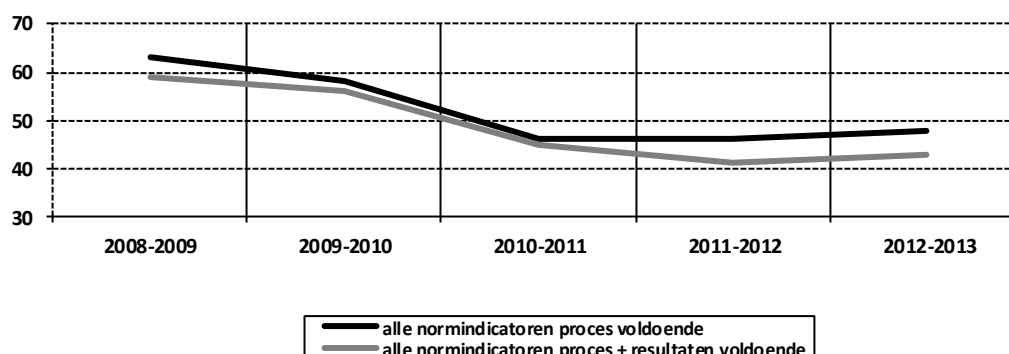
De effectiviteit van het toezicht kan, aldus de Inspectie, ook worden afgemeten aan de snelheid waarin scholen zich ontwikkelen van (zeer) zwak naar basis. In grafiek 3.1 is deze doorlooptijd weergegeven. Hierbij heeft de Inspectie geen onderscheid gemaakt tussen primair en voortgezet onderwijs.

GRAFIEK 3.1: GEMIDDELDE DOORLOOPTIJD IN JAREN VAN (ZEER) ZWAKKE SCHOLEN NAAR BASISARRANGEMENT

Uit de grafiek blijkt dat de gemiddelde doorlooptijd van (zeer) zwakke scholen naar basis in de afgelopen jaren niet is afgenomen. Het gaat hier om een traject van zeer zwak óf zwak naar basis. Dat betekent dat zwakke scholen die eerst zeer zwak worden alvorens weer het basisarrangement te hebben verkregen, samen zijn gevoegd. Als alleen wordt gekeken naar de doorlooptijd waarin zeer zwakke scholen weer het basisarrangement hebben gekregen, is er de laatste jaren wel een dalende lijn waarneembaar. Het gaat hierbij echter om kleine aantallen.⁶

3.4.3 Onderwijsproces

De Inspectie heeft, zoals in hoofdstuk 2 is aangegeven, diverse indicatoren voor de kwaliteit van het onderwijsproces ontwikkeld. Uit grafiek 3.2 blijkt dat er een dalende lijn zichtbaar is van zowel scholen waarbij alle normindicatoren voor het onderwijsproces voldoende zijn als van scholen waarbij niet alleen alle indicatoren van het onderwijsproces maar ook de eind- en tussenresultaten voldoende zijn. Uit de zelfevaluatie blijkt niet op welk(e) schooltypen(n) deze gegevens betrekking hebben.

GRAFIEK 3.2: PERCENTAGE SCHOLEN WAARBIJ ALLE PROCES- EN NORMINDICATOREN ALLE VOLDOENDE ZIJN

Hieruit wordt door de Inspectie de conclusie getrokken dat het toezicht niet effectief is geweest als het gaat om het onderwijsleerproces; er is immers geen stijgende maar een licht dalende lijn te zien in het percentage scholen waarbij alle norm- en procesindicatoren voldoende zijn.

⁶ N zeer zwakke scholen sinds 2009: 35, 58, 40, 28, 18. N zwak plus zeer zwak in die periode: 205, 268, 181, 136, 84.

3.4.4 Bestuursinsteek

Besturen worden in de huidige werkwijze nadrukkelijk aangesproken op hun verantwoordelijkheden voor de kwaliteit van het onderwijs. Door het bestuur (de raad van bestuur of de bestuurder) nadrukkelijk als kwaliteitsbewaker aan te spreken, wordt daar de verantwoordelijkheid gelegd en zal het bestuur minder snel geneigd zijn geen of minimale activiteiten te ondernemen op het terrein van kwaliteitsbewaking en verbetering. De vraag naar de daadwerkelijke effectiviteit van de bestuursinsteek is op zijn inhoudelijke gevolgen niet te toetsen. Er kan, zo stelt de zelfevaluatie, wel gekeken worden naar randvoorwaardelijke effecten en naar percepties vanuit de besturen zelf. Een merkbaar gevolg ervan zou zijn dat de besturen zich veel meer dan tevoren op de hoogte willen stellen van het functioneren van hun scholen. Het leidt tot een sterkere betrokkenheid van de besturen bij de kwaliteitszorg. Vooral besturen met een of meer zeer zwakke scholen, zeggen zich veel bewuster te zijn geworden van hun verantwoordelijkheid op dit punt. Of dat echter ook geldt voor besturen waar zich geen grote problemen hebben voorgedaan, is niet bekend.

3.4.5 Conclusie

Concluderend stelt de zelfevaluatie dat de ontwikkeling van het aantal (zeer) zwakke scholen laat zien dat het risicogerichte toezicht zeer effectief is geweest in het waarborgen van de basiskwaliteit van scholen. Die effectiviteit blijft evenwel beperkt tot de categorie (zeer) zwakke scholen. Scholen die voldoen aan de norm van basiskwaliteit, blijken zich niet verder te verbeteren. Als wordt gekeken naar de doorlooptijd, kan geconcludeerd worden dat het risicogerichte toezicht van de inspectie effectief lijkt om scholen van zeer zwak naar basis te krijgen, maar minder effectief als het gaat om zwakke scholen. De effectiviteit van het toezicht als het gaat om het onderwijsproces lijkt niet groot. De effectiviteit van de bestuursinsteek is niet te toetsen.

4 Evaluatie van het risicogerichte toezichtsmodel

4.1 Inleiding

Nadat hiervoor de toezichtsstrategie van de Inspectie uiteen is gezet en de hoofdpunten van de zelfevaluatie zijn belicht, wordt de toezichtsstrategie van het risicogerichte toezicht in dit hoofdstuk nader belicht en door het externe onderzoeksteam beoordeeld. De evaluatie begint met een algemene beschouwing op het model. Vervolgens komen de centrale thema's van de evaluatie aan de orde, de selectiviteit (par. 4.3), de efficiency (par. 4.5) en de effectiviteit van het risicogerichte toezicht (par. 4.5). Vervolgens wordt aandacht besteed aan de tevredenheid van het onderwijsveld (par. 4.6) en wordt het hoofdstuk afgesloten af met enkele observaties over de interne toepassing van de strategie (par. 4.7) en een conclusie (par. 4.8).

4.2 Het functioneren van het toezichtsmodel

De systematiek van het risicogerichte onderwijstoezicht bestaat sinds 2007 en functioneert inmiddels dus ruim zeven jaar. Aan het toezichtsmodel ligt een heldere en expliciete systematiek ten grondslag die voor het veld goed kenbaar is. Daaruit blijkt een weloverwogen visie op de verdeling van verantwoordelijkheden voor de kwaliteit van het onderwijs. De centrale actor in het model is het schoolbestuur, het aangrijpingspunt voor het onderwijstoezicht. Het schoolbestuur wordt door de Inspectie aangesproken wanneer de kwaliteit van het onderwijs te wensen overlaat. Vanzelfsprekend is het niet het schoolbestuur alleen dat daadwerkelijk verbeteringen door kan voeren. Van grote betekenis daarbij zijn ook de directeur/vestigingsmanager en de leerkracht. Door het bestuur van de school aan te spreken wordt verondersteld dat noodzakelijke veranderingen tot stand komen. Deze bestuursinsteeek wordt door schoolbesturen (h)erkend en ook gevolgd. Het schoolbestuur is uiteindelijk verantwoordelijk en voelt zich ook aangesproken door de oordelen van de Inspectie. Daarbij valt op dat op kritische rapporten van de Inspectie niet wordt gereageerd met weerstand. Veeleer verwelkomen bestuurders de signalen van de Inspectie omdat ze door die signalen worden geattendeerd op een ontwikkeling die aan de aandacht is ontsnapt, omdat met een externe prikkel de mogelijkheid ontstaat intern veranderingen te realiseren of omdat erkenning van het probleem binnen de school op alle lagen (ouders, leerkrachten en bestuur) kan worden omgezet in een eensluidende reactie. Omdat de systematiek van het risicogerichte toezicht transparant en goed vindbaar is en de communicatie door de Inspectie daarover helder, ontstaat betrekkelijk weinig ruis op de lijn.

Dat wil echter niet zeggen dat er geen kritische geluiden in het veld zijn. Zo ligt een mogelijke tekortkoming van het risicogerichte op het vlak van de input van de belanghebbenden op de onderwijskwaliteit. Het model gaat ervan uit dat het schoolbestuur zich extern verantwoordt voor de onderwijskwaliteit aan belanghebbenden en de Inspectie. Belanghebbenden kunnen wensen, klachten en ervaringen inbrengen bij het schoolbestuur en daarmee de onderwijskwaliteit beïnvloeden. Draagt dat onvoldoende bij, dan kunnen die signalen bij de Inspectie terecht komen. Het is om te beginnen de vraag hoe expliciet de externe verantwoording voor de onderwijskwaliteit richting de belanghebbenden precies vorm wordt gegeven. De zelfevaluatie geeft daarover geen uitsluitsel. Aan de gesprekken met bestuurders en schoolleiders ontleen wij twijfel of die verantwoording aan de belanghebbenden altijd voldoende plaatsvindt. Evenmin is duidelijk of schoolbesturen ontvankelijk zijn voor signalen van belanghebbenden en, als dat niet het geval is, of belanghebbenden dan met hun signalen bij de Inspectie terecht kunnen. Of klachten en wensen van ouders, leerlingen, vervolgscholen en werkgevers worden gebruikt als input voor de risicoanalyse van de Inspectie is evenmin bekend. Een andere daarmee verbonden vraag is of berichten in de (social) media over de gang van zaken op scholen door de Inspectie worden benut als input voor de risicoanalyse. Het voornemen dat te doen bestaat wel, maar tot op heden lijkt dat nog niet het geval. Het functioneren van dit deel van het theoretisch model van het onderwijstoezicht zou verduidelijkt kunnen worden.

4.3 Selectiviteit

De essentie van het model van risicogericht onderwijs-toezicht is dat door het toezicht te richten op een selectie van scholen waarvan de onderwijskwaliteit onvoldoende is of dreigt te worden, de belasting van scholen beperkt wordt tot een minimum en de kosten van het toezicht voor de Inspectie eveneens beperkt zullen zijn. De risicoselectie wordt gebaseerd op een analyse van de opbrengsten aan de hand van een aantal indicatoren en daarop gebaseerde beslisregels. De risicoselectie werkt optimaal wanneer het aantal ‘false negatives’, de scholen die ten onrechte als risicovol worden aangemerkt, en het aantal ‘false positives’, de scholen die ten onrechte als risicoloos worden aangemerkt, beperkt blijft. In beide gevallen geschiedt de toets van de risicoanalyse tijdens het schoolbezoek dat volgt op het oordeel ‘ernstig risico’, dan wel tijdens het reguliere schoolbezoek, dat in beginsel op iedere school eens in de vier jaar plaats vindt.

De Inspectie claimt in de zelfevaluatie, zo hebben we in par. 3.2 beschreven, dat de risicoanalyse goed werkt en dat die ook steeds selectiever is geworden. Een vergelijking tussen de risicoanalyse 2007/2008 en die van 2013 laat zien dat het aantal scholen in het PO met een ernstig risico op basis van de opbrengsten, dat toch in een basisarrangement terecht is gekomen, is gedaald. Hier kan op twee manieren worden gekeken. Wordt het aantal scholen met ernstig risico afgezet tegen de totale populatie in de tabellen 3.1 en 3.2, dan is sprake van een afname van 4% (26 van 640) naar ruim 1% (89 van 6991). Wordt het aantal scholen met ernstig risico dat in een basisarrangement is terecht gekomen afgezet tegen het totale aantal scholen met ernstig risico, dan is nog steeds sprake van een daling, zij het dat die gering is (van 51% naar 48,9%).

Een ander punt is de ‘false positives’, de scholen die op basis van de opbrengstenanalyse geen risico leken te hebben, maar die op basis van het reguliere schoolbezoek toch als (zeer) zwak werden gekwalificeerd. Vanuit het oogpunt van bescherming van de leerlingen is dit een nog belangrijker punt. Hier doet de risicoanalyse het in het PO ook goed, gemeten naar het aantal scholen dat ten onrechte niet in de risicoanalyse naar boven is gekomen. Opnieuw is dit cijfer op twee manieren te bepalen. Als deel van het totaal gaat het in de cijfers over 2007/2008 om (26 van 640): 4%. In 2013 betreft het 62 op een totaal van 6991 scholen: 1%. Als deel van de totale populatie ‘(zeer) zwakke’ scholen in de tabellen 3.1 en 3.2 gaat het nog steeds om een daling: van 50% (26 van 51) naar 31% (62 van 197).

De kans op ‘false negatives’ is volgens de cijfers dus gedaald, net als die op ‘false positives’. De risicoanalyse laat overigens tussen 2007/2008 en 2013 een duidelijke afname zien van het aantal scholen met een ernstig risico op basis van de opbrengsten: dat aantal daalt van 7% naar 2%. De Inspectie spreekt van het ‘scherper afstellen’ van de risicoanalyse. Dat lijkt echter een verwarrende term: in werkelijkheid is de analyse juist minder kritisch geworden.

Bij bovenstaande vergelijkingen tussen de tabellen 3.1 en 3.2 moet wel worden opgemerkt dat die strikt genomen moeilijk te maken zijn. De indeling in categorieën verschilt, de normen die zijn aangelegd zijn ook gewijzigd in de loop van de tijd en bovendien zijn – zeker in tabel 3.2 – scholen meegenomen die niet zijn bezocht en die dus vrijwel per definitie in het basisarrangement vallen. In feite zou de conclusie dus moeten zijn dat niet te achterhalen is of de risicoanalyse ‘scherper’ is afgesteld.

De cijfers over het VO laten tussen 2011 en 2013 een toename zien van het aantal ‘false negatives’ van 1,5% (35 van 2442) naar 3% (80 van 2527). De ‘false positives’ blijven stabiel op ruim 1,5% (37 van 2442 in 2011 en 43 van 2527 in 2013). De risicoanalyse in het VO lijkt het dus (iets) minder goed te doen dan in het PO. Ook hiervoor geldt overigens dat veel scholen die wel zijn meegenomen in de tabellen, niet zijn bezocht waardoor ze vrijwel per definitie een basisarrangement hebben gekregen. Om een betere inschatting te kunnen maken van de ‘false positives’ en ‘false negatives’ zouden alleen scholen in de analyse betrokken moeten worden die daadwerkelijk bezocht zijn.

CASUS: IBN GHALDOUN

De gang van zaken rond Ibn Ghaladoun is te zien als een testcase voor de selectiviteit van het toezicht door de Inspectie. De vraag is immers hoe het bij deze school zo uit de hand kon lopen zonder dat de Inspectie maatregelen nam. De afgelopen jaren hebben zich meerdere incidenten op scholen voorgedaan waarbij deze vraag gesteld kan worden.

Rond Ibn Ghaladoun, een brede scholengemeenschap (VWO, HAVO en VMBO) in Rotterdam met vier locaties, waren vanaf 2007/2008 problemen. De Inspectie beoordeelde in 2008 alle afdelingen als zeer zwak, maar in 2010 werd voor alle opleidingen het basisarrangement toegekend. In 2012 werden opnieuw twee afdelingen zwak. In 2013 kwam een diefstal van eindexamens aan het licht. Na onderzoek door de Inspectie, waarin de conclusie werd getrokken dat een verbetering van de onderwijskwaliteit niet binnen twee jaar gerealiseerd zou kunnen worden, werd in het najaar van 2013 de bekostiging van de school gestopt.

In een zelfevaluatie trok de Inspectie de volgende conclusies:

- er was sprake van veel, maar fragmentarisch op de school gericht toezicht, waarbij het aan samenhang ontbrak
- het toezicht is steeds overeenkomstig het toezichtkader en de beslisregels uitgevoerd, maar die gestandaardiseerde werkwijze zorgde er voor dat aan belangrijke aspecten te weinig aandacht werd gegeven
- bij de uitvoering van het toezicht is de context waarin Ibn Ghaladoun opereerde onvoldoende meegewogen
- de toezichthistorie heeft te weinig gewicht gekregen bij de beoordelingen en de noodzaak om tot escalatie over te gaan. Elke bestuurder en schoolleider kreeg het voordeel van de twijfel
- signalen zijn opgepakt, maar te laat

Welke lessen heeft de Inspectie uit de casus 'Ibn Ghaladoun' getrokken? Bij de jaarlijkse risicoanalyse van scholen worden aspecten van kwaliteit, financiën en bestuurlijk handelen meer geïntegreerd gezien. Hierdoor ontstaat meer samenhang in de oordeelsvorming van de Inspectie zodat risico's en tekortkomingen van scholen eerder aan het licht kunnen komen. De inspectie gaat ook op bestuursniveau beoordelen of een onderzoek naar het bestuurlijk handelen gewenst is. Hierbij zijn geconstateerde risico's op het gebied van kwaliteit en financiën leidend.

De cijfers overziend kan de conclusie zijn dat de methode van risicogericht toezicht er in lijkt te slagen de groep scholen te identificeren die voor een aangepast toezichtsarrangement in aanmerking komt, zonder dat een groot aantal scholen ten onrechte als risicovol wordt gekwalificeerd en zonder dat al te veel scholen ten onrechte als risicoloos worden aangemerkt. Het percentage 'false negatives' en 'false positives' is zowel voor PO als VO beperkt: het gaat om een zeer gering deel van de totale populatie, afhankelijk van de categorie 'false negatives' of 'false positives' en het soort onderwijs 1%-3% in 2013.

Een kanttekening bij bovenstaande beschouwing van 'false negatives' is wel dat hierbij gekeken is naar de scholen die volgens de risicoanalyse een ernstig risico hebben. Dat oordeel wordt toegekend aan scholen waarvan de opbrengsten voor het derde achtereenvolgende jaar onvoldoende zijn. Dat na het schoolbezoek blijkt dat slechts een beperkt deel van die scholen alsnog in het basisarrangement terecht komt is dan ook niet verwonderlijk. Het gaat in die gevallen immers om uitzonderingssituaties: veelal kleine scholen, een situatie van zij-instroom etc. Een nog belangrijker kanttekening bij de in de zelfevaluatie gepresenteerde cijfers is dat de ontwikkeling in de tijd met de nodige onzekerheden is omgeven. Zoals in paragraaf 2.3 is beschreven, is het toezichtskader in de onderzochte periode inhoudelijk aangepast, is er sprake van niet of slechts moeilijk vergelijkbare categorieën en zitten er in de cijfers veel scholen die niet zijn bezocht.

Bij een bespreking van de mate van selectiviteit van het onderwijstoezicht hoort ook een opmerking over de gestelde ondergrens voor kwaliteit. Het model van risicogericht toezicht slaagt er goed in de scholen met problemen in hun onderwijskwaliteit te selecteren op basis van een risicoanalyse van de opbrengsten. Daarmee is echter nog niet gezegd dat de juiste maatstaf wordt gehanteerd voor kwaliteit. Er zijn signalen die er op wijzen dat de voor kwaliteit gehanteerde maatstaf niet zeer streng is. Inspecteurs stellen in dit verband wel dat er heel wat mis moet zijn op een school voordat aan de onderwijskwaliteit het oordeel zeer zwak wordt toegekend. Dat neemt niet weg dat gegeven de gedefinieerde normen voor kwaliteit het risicogerichte toezicht er redelijk goed in lijkt te slagen de ondermaats presterende scholen te selecteren.

4.4 Efficiency

Bij efficiency gaat het om een onderwerp dat in het verlengde van selectiviteit ligt. Immers, hoe selectiever de risicoanalyse, hoe meer de Inspectie er in kan slagen de inzet te richten op de onderkant van de scholen die onvoldoende onderwijskwaliteit laten zien. De Inspectie claimt in de zelfevaluatie dat het risicogerichte toezicht efficiënt is. Die claim is in algemene zin te baseren op de systematiek van het risicogerichte toezicht: in plaats van elke school elk jaar te bezoeken wordt jaarlijks een risicoanalyse op basis van de opbrengsten uitgevoerd. Alleen die scholen met een 'ernstig risico' worden vervolgens bezocht. Daarnaast wordt net als voorheen elke school ten minste eens in de vier jaar bezocht. Het risicogerichte toezicht is dus efficiënter wanneer de kosten voor de Inspectie van het schoolbezoek van scholen die als risicovol zijn aangemerkt geringer zijn dan de kosten die voorheen gemoeid waren met de jaarlijkse bezoeken. Daarvan lijkt inderdaad sprake. De Inspectie heeft meer capaciteit in de analyse van de cijfers moeten investeren, maar de kosten van schoolbezoeken nemen af. Die kosten, in termen van de tijd die is gemoeid met een schoolbezoek aan een school met onvoldoende kwaliteit, nemen weliswaar toe, maar doordat daartegenover een sterke daling staat van de kosten van de reguliere schoolbezoeken (eens per vier jaar), is per saldo toch van een grotere efficiency sprake. De zelfevaluatie laat dat ook zien. Uit de cijfers van de Inspectie blijkt dat sprake is van minder fte's bij de Inspectie, maar tegelijkertijd slaagt de Inspectie er met die geringere capaciteit in die schoolbezoeken uit te voeren die geïndiceerd zijn op grond van de risicoanalyse.

Een andere vraag is hoe de toezichtslasten zich voor de scholen ontwikkelen. De Inspectie spreekt in de ex ante evaluatie van het risicogerichte toezicht van een beoogde halvering van de toezichtslasten. Duidelijk is dat de introductie van de systematiek er voor heeft gezorgd dat de reguliere schoolbezoeken niet langer elk jaar, maar eens in de vier jaar worden bezocht. De 'goede' scholen worden dus minder belast. Uit de cijfers blijkt dat de schoolbezoeken die volgen op het oordeel 'ernstig risico' veel intensiever zijn dan het jaarlijkse schoolbezoek voorafgaand aan de introductie van het risicogerichte toezicht. De toezichtslasten voor de (zeer) zwakke scholen zijn dus duidelijk toegenomen, maar omdat het gaat om een (zeer) klein deel van de totale populatie zijn per saldo de totale toezichtslasten voor de scholen afgenomen. Deze cijfers gelden voor het PO, voor het VO zijn die niet beschikbaar. Dat zich daar eenzelfde ontwikkeling heeft voorgedaan is niet onaannemelijk.

Behalve door de schoolbezoeken worden scholen ook belast door de verzoeken om informatie die de Inspectie tot ze richt. Deze bevragingslast is door de Inspectie in kaart gebracht. Uit de zelfevaluatie blijkt dat die nagenoeg gelijk is gebleven voor het PO en licht gedaald is voor het VO. Het cijfer voor de bevragingslast representeert het aantal vragenlijsten dat bij de scholen is uitgezet. Daarmee is nog geen specificatie beschikbaar voor de tijd die elke afzonderlijke vragenlijst kost. Een sluitende analyse hiervan is moeilijk te maken. Op basis van de gegevens van de Inspectie staat niet vast hoeveel tijd daadwerkelijk met de bevraging is gemoeid. Jaarlijks zijn er andere vragenlijsten en de met de vragenlijsten benodigde invultijd wordt op een steeds andere wijze ingeschat. Zo wordt de invultijd in 2010 en 2011 door de Inspectie ingeschat, terwijl de invultijd in 2012 en 2013 door de scholen diende te worden aangegeven. De representativiteit van de (gemiddelde) door scholen aangegeven invultijd is ook twijfelachtig.

Zo blijkt dat in de laatste twee jaar de minimale en maximale invultijd soms enorm uit elkaar loopt. Zo bedroeg de invultijd die een school rapporteerde voor een onderzoek naar Onderwijstijd 2.800 minuten, tegenover 10 minuten die het een andere school zou hebben gekost. De inschatting van de Inspectie is dat de daadwerkelijke tijd die scholen nodig hebben voor het invullen van vragenlijsten, afneemt (onder meer door kortere vragenlijsten), maar we hebben dit niet kunnen objectiveren. Duidelijk is wel dat het aantal bevragingen, waar de Inspectie in de zelfevaluatie onder meer op baseert dat het toezicht voor scholen efficiënter is geworden, geen goede indicator voor efficiency vormt.

Desondanks kan de conclusie getrokken worden dat de efficiency van het onderwijstoezicht zich positief lijkt te ontwikkelen. Met minder mensen is de Inspectie er in geslaagd haar inzet te concentreren op die scholen die dat het meest nodig hebben: de scholen waarbij volgens de risicoanalyse sprake zou zijn van een ernstig risico. Voor die scholen zijn de toezichtslasten duidelijk toegenomen, voor de scholen die voldoende scores is sprake van een duidelijke daling van de belasting door de Inspectie. Die verdeling lijkt in lijn met de grondslag van het risicogerichte toezicht, die zich vooral richt op kwaliteitsverbetering van de (zeer) zwakke scholen.

4.5 Effectiviteit

Uit de zelfevaluatie blijkt dat de Inspectie van mening is dat haar toezicht effectief is als het gaat om het waarborgen van de basiskwaliteit van scholen. Daar zijn de toezichtsactiviteiten van de Inspectie ook primair op gericht. Als we ervan uitgaan dat de focus op opbrengsten terecht is, oftewel, dat de normen op grond waarvan een school zwak wordt de juiste zijn, kan dat voor het primair onderwijs inderdaad geconcludeerd worden. Het aantal zwakke scholen binnen deze sector is de afgelopen jaren substantieel gedaald. Het is aannemelijk dat deze daling aan de inspanningen van de Inspectie te danken is, hoewel er wellicht ook andere oorzaken zouden kunnen zijn. Zo zijn er de laatste jaren vooral in krimpgebieden kleine scholen gesloten die relatief vaak zwak waren. Dit verklaart niet de (volledige) daling van het aantal zwakke scholen in het primair onderwijs, maar dergelijke deels verklarende factoren zijn door de Inspectie niet in de zelfevaluatie aan de orde gesteld. De daling van het aantal zwakke scholen wordt daarin volledig toegeschreven aan het toezicht.

Voor het voortgezet onderwijs is de daling van het aantal (zeer) zwakke scholen beperkt. Dat betekent niet dat het toezicht niet effectief is, het is immers bijvoorbeeld niet duidelijk hoeveel scholen in het voortgezet onderwijs (zeer) zwak waren geweest zonder het systeem van risicogericht toezicht. Als gekeken wordt naar de ontwikkeling van de normen die zich richten op het onderwijsproces, dan is er geen positieve tendens zichtbaar. Ook hiervoor geldt dat niet duidelijk is hoe dit zich ontwikkeld zou hebben zonder toezicht door de Inspectie. Bovendien kan door een scherpere focus van de Inspectie op het onderwijsproces (hetgeen bestaat uit grotendeels kwalitatieve normen) er ook voor gezorgd hebben dat meer scholen hier negatief op scoren.

Het is zonneklaar dat het aantal (zeer) zwakke scholen in het PO de afgelopen (substantieel) is gedaald. Het causale verband met het (risicogerichte) toezicht door de Inspectie is niet onomstotelijk aantoonbaar, maar lijkt aannemelijk. Lukt het de geïdentificeerde (zeer) zwakke scholen ook snel weer op een voldoende niveau terecht te komen? De zelfevaluatie laat zien dat het verbeterproces voor de zeer zwakke scholen de afgelopen jaren (vergelijking 2009/2010 met 2013/2014) is versneld van 2,5 tot 3 jaar naar 1,5 naar 2 jaar. De zwakke scholen doen er iets langer over dan voorheen, van iets sneller dan 1,5 naar iets langer dan 1,5 jaar. Over de hele linie valt op dat scholen die als (zeer) zwak worden gekwalificeerd toch betrekkelijk snel weer op voldoende niveau terecht kunnen komen.

Een belangrijke kanttekening die bij het voorgaande en bij de effectiviteit van het onderwijs-toezicht kan worden gemaakt is dat hierbij de focus ligt op de (zeer) zwakke scholen. Ook bij scholen die niet (zeer) zwak zijn is de onderwijskwaliteit uiteraard van belang. Of het toezicht

een gunstige invloed heeft op de onderwijskwaliteit in brede zin, kan echter niet worden vastgesteld. De zelfevaluatie biedt daarvoor ook onvoldoende informatie. Dat scholen op voldoende niveau scoren in de risicoanalyse wil niet zeggen dat zij niet beter zouden kunnen. Juist om die reden is de Inspectie bezig met de ontwikkeling van Toezicht2020, dat zich richt op 'stimulerend toezicht' in de breedte van het onderwijs.

Op dit moment gaat van het toezicht dat de Inspectie uitoefent niet of nauwelijks een kwaliteitsimpuls uit die gericht is op scholen met een basisarrangement. De conclusie is dan ook dat het toezicht van de Inspectie slechts een beperkte invloed heeft op de algehele ontwikkeling van de kwaliteit van het onderwijs. Het signaal van de Inspectie voor scholen waarvan de onderwijskwaliteit voldoende is, is dat de Inspectie hen met rust laat. Dat kan voor die scholen goed zijn en ze de ruimte geven zich op eigen kracht te ontwikkelen. Het is echter ook mogelijk dat scholen die net voldoende scoren, op dat niveau blijven hangen; het onderwijstoezicht is voor hen geen prikkel zich te verbeteren. De Inspectie zou hierover de komende jaren informatie moeten zien te verzamelen, in aanloop naar de introductie van Toezicht2020, zodat onomstotelijk vast staat wat de invloed is van het toezicht op de kwaliteit van het onderwijs, ook wanneer de onderwijskwaliteit op zich voldoende is.

Overigens geven opbrengsten niet altijd de werkelijke meerwaarde van een school voor kinderen weer. Een gemiddelde CITO-score van 535 of een cijfer 6,5 op een centraal eindexamen kan voor de ene school een bovengemiddelde prestatie zijn terwijl het voor een andere school lager is dan verwacht mag worden. Weliswaar worden resultaten gewogen aan de hand van het opleidingsniveau van de ouders, maar er zijn meer factoren die de 'uitgangssituatie' van een leerling bepalen. IQ, sociaal-economische achtergrond, bevolkingsopbouw, een stimulerende thuissituatie, kennis van de Nederlandse taal, grootte van klassen (en de invloed op extremen op het gemiddelde), het zijn factoren die een gemiddelde CITO-score beïnvloeden maar die niet (volledig) verdisconteerd zijn in de generieke wegingsfactor, die uitsluitend het opleidingsniveau van de ouders meeweegt.⁷ Bovendien zijn de marges binnen de wegingsfactoren ruim; er zit een verschil tussen ouders die gemiddeld bijvoorbeeld aan de onderkant van MBO-niveau zitten en die aan de bovenkant zitten, maar er wordt dezelfde wegingsfactor op toegepast.

4.6 Neveneffecten

De vraag is of de door de Inspectie gehanteerde systematiek ook (ongewenste) neveneffecten kent. Dat lijkt zeker het geval. Er is bij de Inspectie een focus op opbrengsten, omdat een school bij onvoldoende opbrengsten in drie opeenvolgende jaren vrijwel per definitie (zeer) zwak wordt en de opbrengsten de basis vormen voor de risico-analyses. Door te focussen op opbrengsten ontstaat bij scholen de neiging zich in het onderwijs sterk op de opbrengsten te concentreren. Jaarlijks zijn er slechts enkele gevallen waarin het onderwijsproces dermate positief werd beoordeeld dat het de onvoldoende opbrengsten compenseerde en de school alsnog in het basisarrangement terecht kwam. Scholen vertonen daardoor de neiging de opbrengsten te overaccentueren, zeker als de school al in de gevarenzone zit en voldoende opbrengsten op een gegeven moment noodzakelijk zijn om niet zwak te worden verklaard. Hierdoor kunnen perverse prikkels optreden. Scholen voegen zich naar de normen van de Inspectie. Ze richten zich helemaal op het behalen van voldoende opbrengsten, hetgeen ten koste kan gaan van andere activiteiten.

Daarnaast zijn veel betrokkenen (schoolbesturen, schooldirecteuren, inspecteurs) het erover eens dat opbrengsten te manipuleren zijn. In het PO werd voorheen wel getracht de resultaten de zwakkere leerlingen buiten beschouwing te laten. Nadat de Inspectie daartegen is gaan optreden (door de CITO-score 517 toe te kennen aan leerlingen die zonder onderbouwing buiten de scores zijn gebleven) is een neveneffect dat tegenwoordig breed optreedt, zeker bij (zeer)

⁷ Wel geldt sinds augustus 2014 een nieuwe werkinstructie voor het herberekenen van de eindresultaten. Daarbij kunnen leerlingen buiten de beoordeling worden gelaten wanneer ze toelaatbaar zijn tot het praktijkonderwijs of het voortgezet speciaal onderwijs, wanneer leerlingen in groep 7 of 8 zijn ingestroomd en wanneer leerlingen sinds kort in Nederland verblijven en om die reden het Nederlands minder goed beheersen.

zwakke scholen, het intensief oefenen met toetsen. Dat hoeft geen bezwaar te zijn, maar is vooral dan een ongewenst effect wanneer daardoor onvoldoende tijd over blijft voor zaken die overigens de onderwijskwaliteit bepalen, maar die niet direct in de opbrengsten terug komen. Bovendien kan het hierdoor voorkomen dat leerlingen een hogere dan reële CITO-score behalen. Hier kunnen scholen in het VO weer last van krijgen, omdat ze wellicht naar een hoger schooltype gaan dan reëel is en ze dan de doorstroomcijfers in de onderbouw (één van de opbrengstennormen voor het VO) negatief beïnvloeden.

Scholen in het VO kunnen proberen te voorkomen dat leerlingen zich inschrijven voor een hoger schooltype dan op grond van de CITO-score en/of het lerarenadvies geïndiceerd was, omdat ze anders de doorstroming in de onderbouw negatief beïnvloeden. In algemene zin geldt dat het voor scholen in het VO die willen sturen op opbrengsten, in principe voordelig is dat een leerling op een zo laag mogelijk onderwijsniveau instroomt. Een dergelijke leerling zal dan relatief goed presteren en beïnvloedt de opbrengsten van de school positief. Er zijn scholen, zo is ons in interviews gebleken, die hierop proberen te sturen. De Inspectie bevraagt VO-scholen overigens indien er opvallende verschillen zijn tussen de adviezen die een leerling van PO-school krijgen en de daadwerkelijke plaatsing op een school.

In de bovenbouw van het VO kan een soortgelijke prikkel ontstaan als in het PO, namelijk dat intensief geoefend wordt (“gedrild wordt”, aldus een gesprekspartner) op het centraal eindexamen. Volgens een geïnterviewde schoolbestuurder wordt er “gesjoemeld bij het leven” om bijvoorbeeld het onderbouwendement op te krikken. Het voert in het kader van deze evaluatie te ver om hier diep op in te gaan, maar duidelijk is dat de focus van de Inspectie op opbrengsten met zich meebrengt dat scholen er soms alles aan proberen te doen om de opbrengsten voldoende te laten zijn. In het primair onderwijs is frauduleus handelen zelfs niet geheel te voorkomen. Inspecteurs kennen voorbeelden van scholen waar leerkrachten CITO-toetsen beïnvloeden door aanwijzingen te geven aan leerlingen, ze extra tijd te geven of waar de gelegenheid wordt geboden toetsen opnieuw te maken of achteraf antwoorden aan te passen. Inspecteurs komen ook op scholen waar tijdens de toets posters met maten aan de muur hangen. Indien een school opeens een opvallend hoge gemiddelde CITO-score behaalt wordt daar vaak onderzoek naar gedaan door de Inspectie, maar ook als er een vermoeden van fraude of manipulatie is, is dat vaak niet hard te maken.

4.7 Tevredenheid

In de media duiken regelmatig kritische geluiden op over het werk van de Inspectie. De Inspectie zou scholen overladen met administratieve rompslomp en toetsen aan irrelevante normen, zo luidt de kritiek vaak. Opvallend is dat deze ontevredenheid niet gestaafd wordt in de tevredenheidsonderzoeken die de Inspectie onder schoolbesturen houdt. Daaruit blijkt dat scholen over het algemeen zeer tevreden zijn over het inspectiebezoek en de daaraan gekoppelde verbeterafspraken en dat deze tevredenheid zelfs toeneemt. Scholen zijn vooral tevreden over het contact met de inspecteur, het bezoek zelf en het resultaat van het bezoek. Ook uit de door ons gevoerde gesprekken blijkt in grote lijn tevredenheid over de inspectie, juist ook bij zwakke scholen. Kritiek van de Inspectie wordt erkend en de kwalificatie ‘(zeer) zwakke school’ kan gebruikt worden om verbeteringen door te voeren en het belang daarvan in brede kringen (besturen, directie, leraren, ouders) onderkend te krijgen.

4.8 Inspectie van het Onderwijs intern

Bij het bestuderen van de wijze waarop de Inspectie het risicogerichte toezicht uitvoert valt op dat sprake is van een binnen de verschillende onderdelen van de organisatie weinig geüniformeerde wijze van werken. Het toezicht op het primair onderwijs en het voortgezet onderwijs zijn organisatorisch en personeel strikt van elkaar gescheiden. Dat leidt tot een merkwaardige consequentie: waar de inspecteurs scholen aanmoedigen van elkaar te leren (‘ga eens kijken bij de scholen waar de leerlingen vandaan komen’, of ‘ga eens kijken naar de

vervolgopleidingen waar de leerlingen naar toe gaan'), is er intern geen sterke traditie van onderlinge uitwisseling. Enige tijd geleden was nog wel sprake van een beleid dat zich richtte op regelmatige uitwisseling van medewerkers, maar momenteel is vooral sprake van specialisatie. Dat leidt tot verschillen in werkwijze, methodiek en cultuur. Naar verluidt is er ook sprake van verschillen in gestrengheid in het land. De verschillen in methodiek komen ook tot uitdrukking in de uiteenlopende kaders. De toezichtskaders zijn goed vindbaar, maar door de voortdurende wijzigingen die daarin worden aangebracht, maakt het geheel niet altijd een goed toegankelijke indruk. Aanvankelijk was sprake van geïntegreerde kaders voor PO en VO, sinds 2013 zijn er afzonderlijke kaders. Wat ook niet helpt is dat de terminologie niet consistent is; de ene keer heet het waarderingskader ook wel kernkader, dan weer is het kernkader een specifiek deel van de indicatoren binnen het waarderingskader. Wat een en ander nog lastiger te doorgronden maakt is dat de beslisregels niet (steeds) zijn opgenomen in de toezichtskaders, maar elders moeten worden gevonden. Het waarderingskader bestaat de ene keer uit vijf domeinen, dan weer uit drie domeinen. De wijze waarop de cijfers worden gepresenteerd (en verzameld?) wijzigt ook voortdurend, waarbij andere uitgangspunten worden gehanteerd en een afwijkende indeling in categorieën aan de orde is. De vergelijkbaarheid in de tijd is daardoor soms een probleem waar ook in dit onderzoek tegenaan is gelopen.

Om verschillen tussen inspecteurs te voorkomen wordt intern veel over casuïstiek gesproken, worden over 'twijfelgevallen' overlegd, worden de teams die op schoolbezoeken gaan wisselend samengesteld en is er overleg tussen inspecteurs over hun waarnemingen tijdens de schoolbezoeken. Die uitwisseling is van groot belang om interbeoordelaarsbetrouwbaarheid te bevorderen, hoewel die nooit kan worden gegarandeerd.

Een punt van kritiek dat de methodiek van het onderwijstoezicht betreft heeft betrekking op de werkwijze waarbij schoolbezoeken tevoren worden aangekondigd. Uit de gesprekken met zowel inspecteurs als schoolbesturen komt naar voren dat dit er toe leidt dat scholen zich op die bezoeken voorbereiden. Zo worden leraren extra gewezen op de eisen die de inspectie stelt aan bijvoorbeeld de opbouw en inhoud van een les of blijven mindere leraren soms een dagje thuis. Het is dan ook de vraag of het beeld dat tijdens de schoolbezoeken bij de inspecteurs ontstaat van het onderwijsproces en de kwaliteitszorg representatief is voor de algemene gang van zaken binnen de school. Uiteraard is een geoefend inspecteur in staat door de werkelijkheid heen te kijken, maar een element van vertekening blijft altijd bestaan.

4.8 Algemeen: concluderend

Onze algemene conclusie is dat het risicogerichte toezicht door de Inspectie er op een efficiënte wijze aan bijdraagt de onderwijskwaliteit op scholen op voldoende niveau te brengen. Het aantal (zeer) zwakke scholen is de afgelopen jaren gedaald en de tijd die het scholen kost van een aangepast arrangement in het basisarrangement terecht te komen is korter geworden. Dat wil echter niet zeggen dat geen kanttekeningen te plaatsen zijn bij de systematiek van risicogericht toezicht.

De eerste kanttekening betreft de focus op opbrengsten. De toespitsing op opbrengsten in de risicoanalyse leidt tot een efficiënte inzet van menskracht bij de Inspectie en tot relatief lage toezichtslasten voor de scholen. Tegelijkertijd kan de focus op de opbrengsten ten koste gaan van het onderwijsproces voor de onderwijskwaliteit. Dat is principieel het geval omdat de analyse op opbrengsten de intensiteit van het toezicht bepaalt, maar is ook vervolgens aan de orde omdat schoolbezoeken altijd in het teken staan van de tekortschietende opbrengsten.

De tweede kanttekening is dat door de focus op opbrengsten het ongewenste neveneffect optreedt dat het voor scholen aantrekkelijker wordt de resultaten te manipuleren. Er zijn veel signalen die er op wijzen dat daarvan inderdaad sprake is.

De derde kanttekening betreft de hoogte van de normen die worden aangelegd. Er wordt wel gesuggereerd dat de Inspectie 'te soft' zou zijn. Dat zou onder meer tot uitdrukking komen in de

tijd die het scholen die als (zeer) zwak zijn gekwalificeerd kost weer op voldoende niveau te komen: dat lukt over het algemeen betrekkelijk snel (gemiddeld 1,5 tot 2 jaar).

De vierde kanttekening betreft de focus van het toezicht op de zwakke en zeer zwakke scholen. Die focus leidt er toe dat het overgrote deel van de scholen in ons land op voldoende niveau terecht zijn gekomen. De ontwikkeling van de onderwijskwaliteit op de andere scholen onttrekt zich echter voor een belangrijk deel aan de waarneming. Het toezicht kan aan de verbetering van die onderwijskwaliteit dan ook niet of nauwelijks een bijdrage leveren.

De vijfde kanttekening is dat de opbrengsten in het PO slechts één generieke wegingsfactor hebben, namelijk het opleidingsniveau van de ouders. Er zijn echter meer factoren aan te duiden die de gemiddelde hoogte van een CITO-toets beïnvloeden.

De algemene conclusie is dus positief voor het risicogerichte toezicht. Ondertussen is sprake van doorontwikkeling van het risicogerichte onderwijstoezicht in de richting van ‘stimulerend’ onderwijstoezicht. Daarmee zou het probleem van de focus op de onderkant zijn verholpen. De vraag is tegelijkertijd hoe dat stimulerend toezicht vormgegeven kan worden. Daarover meer in het volgende hoofdstuk.

5 Conclusie

In deze evaluatie is het risicogerichte toezicht onder de loep genomen. De centrale vraag in deze evaluatie richt zich op het succes van het risicogerichte toezicht in termen van selectiviteit, efficiency en kwaliteit. Hierna wordt deze vraag samenvattend beantwoord.

Het risicogerichte toezicht slaagt er in de risicoanalyse goed in de scholen met (zeer) zwakke onderwijskwaliteit te achterhalen. De systematiek is niet waterdicht: er zijn ‘false positives’, maar het aantal scholen dat het betreft is beperkt en is acceptabel. Het aantal ‘false negatives’ is iets groter. Dat is ook onontkoombaar, tenzij de risicoanalyse een zeer ruime tolerantiegrens zou krijgen. De vraag of het onderwijstoezicht ook effectief is, is uiteraard lastig te beantwoorden. Alternatieve verklaringen voor de impact van het risicogerichte toezicht kunnen niet uitgeschakeld worden. Toch is niet onaannemelijk dat het toezicht een belangrijke bijdrage heeft geleverd aan het verminderen van het aantal (zeer) zwakke scholen in ons land. Scholen doen er gemiddeld 1,5 jaar over van het aangepaste toezichtsarrangement ‘zwak’ weer in het basisarrangement terecht te komen. Zeer zwakke scholen doen daar uiteraard wat langer over. Dat neemt niet weg dat het opvallend is dat scholen gemiddeld relatief (1,5 tot 2 jaar) snel weer in het basisarrangement terecht komen. Het beeld bestaat dat de onderwijskwaliteit van scholen relatief snel onder de kritische grens terecht kan komen, maar ook relatief snel weer op niveau kan geraken. Dat onderstreept het belang van een systematiek waarbij voortdurend, dat wil zeggen van jaar tot jaar, de ontwikkeling van de onderwijskwaliteit op scholen wordt gevolgd.

Een bijzonder kenmerk van het risicogerichte toezicht is dat het vrijwel geheel de nadruk legt op de scholen aan de onderkant van het spectrum. Wanneer een school voldoende scoort komt die in het basisarrangement en is er niet of nauwelijks contact met de toezichthouder. Die komt volgens het algemene uitgangspunt eens in de vier jaar wel een kijkje nemen, maar een intensief schoolbezoek is dat niet. Op aandrang van de Tweede Kamer en ook op initiatief van de Inspectie gaat dat mogelijk veranderen. Toezicht2020 stelt zich ten doel het toezicht ook voor scholen die voldoende presteren te activeren. Dat betekent dat door ‘stimulerend toezicht’ of ‘gedifferentieerd toezicht’ toe te passen ook verbetering van voldoende scholen richting excellentie wordt bevorderd.

De vraag is hoe de Inspectie kan bevorderen dat scholen in het basisarrangement het beter gaan doen. De Inspectie lijkt hier bijna onvermijdelijk op twee gedachten te hinken. Enerzijds is aan het wezen van toezicht op de naleving verbonden dat de Inspectie geen inhoudelijke adviezen kan geven voor de door het schoolbestuur te nemen acties, anderzijds moeten scholen die in het basisarrangement zitten toch aangezet worden tot verbetering van de onderwijskwaliteit. Een van de instrumenten die dan ingezet kan worden is de verdere differentiatie van de onderwijskwaliteit in kwaliteitscategorieën. Waar nu (zeer) zwak en voldoende de onderscheidende grootheden zijn, roept de weg naar ‘gedifferentieerd toezicht’ de behoefte aan verdere specialisatie van de oordelen op. In het gedifferentieerde toezicht zullen scholen die duidelijkheid ook wensen. Deze weg kent risico’s. De verdere detaillering van de waarderingskaders en beslisregels kan leiden tot ongewenste jurisdisering: regulering die steeds verdere verfijning van normen oproept. En dat roept weer het risico op van het ongewenste neveneffect van inzet op de indicatoren die bepalend zijn in de categorisering.⁸ De Inspectie stelt hierover dat de insteek aan de ‘bovenkant’ een wezenlijk andere moet zijn dan die aan de ‘onderkant’. Gedacht moet worden aan dialoog, die tot stand komt tussen professionals. Hoe dat precies vorm moet krijgen is onderwerp van een zoektocht, waarin de Inspectie zich momenteel bevindt. Het is verstandig voor die experimenteerfase de tijd te nemen tegen de achtergrond van de hierboven geschetste risico’s van gedifferentieerd toezicht.

De Inspectie is in haar zelfevaluatie zeer positief over de efficiency van het toezicht, positief over de risico-analyses en gematigd positief over de effectiviteit van het toezicht. In onze analyses hebben we hier kanttekeningen bij gezet, de zelfevaluatie was naar ons oordeel soms iets te

⁸ Zie ook Ferdinand Mertens, Naar gedifferentieerd toezicht: elke school in een vakje! Mei 2014.

positief en eendimensionaal. Dit laat echter onverlet dat – als eindoordeel – ook door ons de drie aspecten alles bij elkaar opgeteld positief gewaardeerd worden.

Bijlage 1: Geraadpleegde documenten

Documenten

CDA, PvdA en ChristenUnie (2007), Coalitieakkoord.
Gemeente Amsterdam (2009), Kwaliteitsaanpak Basisonderwijs Amsterdam.
Inspectie van het Onderwijs, Onderwijsverslagen 2006 t/m 2013.
Inspectie van het Onderwijs, Jaarverslagen 2007 t/m 2013.
Inspectie van het Onderwijs, Jaarwerkplannen 2007 t/m 2014.
Inspectie van het Onderwijs (2006) Onderzoek zeer zwakke scholen.
Inspectie van het Onderwijs (2007) De WOT verwerkelijkt. Nieuw toezicht voor de toekomst.
Inspectie van het Onderwijs (2008) Beslisschema voor het bepalen van een aangepast arrangement (zwakke of zeer zwakke school).
Inspectie van het Onderwijs (2008) Implementatieplan taakstelling, mei 2008.
Inspectie van het Onderwijs (2009) Toezichtkader PO VO 2009. Utrecht, juli 2009.
Inspectie van het Onderwijs (2009) Toezichtkader BVE 2009.
Inspectie van het Onderwijs (2009) Aanvulling voor 2010 op het toezichtkader BVE 2009.
Inspectie van het Onderwijs (2009) Beslisschema's aangepast arrangement PO.
Inspectie van het Onderwijs (2010) Evaluatie Risicogericht Toezicht.
Inspectie van het Onderwijs (2010) Risicoanalyse: Risico's op scholen in beeld.
Inspectie van het Onderwijs (2011) Toezicht op zwakke en zeer zwakke scholen: zo werkt het.
Inspectie van het Onderwijs (2012) De beoordeling van opbrengsten in het basisonderwijs.
Inspectie van het Onderwijs (2012) Ontwikkeling bevragsingslast
Inspectie van het Onderwijs (2012) Tevredenheidsonderzoek besturen primair onderwijsschooljaar 2011/2012.
Inspectie van het Onderwijs (2012) Toezichtkader PO/VO 2012.
Inspectie van het Onderwijs (2013) Analyse en waarderings van opbrengsten PO.
Inspectie van het Onderwijs (2013) Bevragsingslast: stand van zaken 2012.
Inspectie van het Onderwijs (2013) De kwaliteit van het bestuurlijk handelen in het funderend onderwijs.
Inspectie van het Onderwijs (2013) Lessen uit het toezicht op ISG Ibn Ghaldoun.
Inspectie van het Onderwijs (2013) Onderzoek bestuurlijk handelen Islamitische Scholengemeenschap Ibn Ghaldoun.
Inspectie van het Onderwijs (2013) Toezichtkader VO 2013.
Inspectie van het Onderwijs (2013) Uitgangspunten werkwijze waarderingskader.
Inspectie van het Onderwijs (2013) Waarderingskader Voortgezet Onderwijs 2013.
Inspectie van het Onderwijs (2014) Bevragsingslast: stand van zaken 2013.
Inspectie van het Onderwijs (2014) Evaluatie toezicht 2013: het toezicht van PO in getallen.
Inspectie van het Onderwijs (2014) Kwaliteit nader beschouwd alle onderwijssoorten.
Inspectie van het Onderwijs (2014) Meerjaren opbrengsten VO 2014: toelichting.
Inspectie van het Onderwijs (2014) Onderzoek bestuurlijk handelen Stichting voor evangelische scholen (sVES).
Inspectie van het Onderwijs (2014) Oordelen OV 2013 – Kwaliteit meerjarenoverzicht.
Inspectie van het Onderwijs (2014) Resultaten tevredenheidsonderzoek Inspectie van het Onderwijs: Schooljaar 2012/2013.
Inspectie van het Onderwijs (2014) Resultaten tevredenheidsonderzoek Inspectie van het Onderwijs 2012/2013.
Inspectie van het Onderwijs (2014) Tevredenheidsonderzoek Expertisecentra schooljaar 2012/2013.
Inspectie van het Onderwijs (2014) Tevredenheidsonderzoek Primair Onderwijs schooljaar 2012/2013.
Inspectie van het Onderwijs (2014) Tevredenheidsonderzoek Voortgezet Onderwijs schooljaar 2012/2013.
J/M voor ouders (2009) Onderzoek vertrouwen in onderwijs onder ouders.
J/M voor ouders (2012) Onderzoek vertrouwen in onderwijs onder ouders.
Ministerie van OCW (2006), Kaderstellende visie op toezicht: toezicht in vertrouwen, vertrouwen in toezicht.

Ministerie van OCW (2007) Evaluatie Wet op het Onderwijstoezicht (uitgevoerd door EIM).

Literatuur

- ABDTopconsult, *Externe validering Lessons Learned Inspectie van het Onderwijs: Een analyse over de effectiviteit van het toezicht op Ibn Ghaldoun*, Den Haag: 2014.
- Beek, S & Van den Heuvel, P., *Het stoplicht sprong op rood! Verslag van onderzoek naar bestuurlijke interventies naar zeer zwakke scholen*. Den Bosch: KPC groep 2007.
- Matthews, P & Sammons, P., *Improvement through inspection: an evaluation of the impact of Ofsted's work*. London: Institute of Education, University of London/Ofsted: 2004.
- F. Mertens, *Naar gedifferentieerd toezicht: elke school in een vakje!*, 2014.
- Oberon, *evaluatie van het toezichtkader vo 2013*, Utrecht: 2014.
- Onderwijsraad, *Doortastend onderwijstoezicht*. Den Haag: 2006
- Onderwijsraad, *Onderwijspolitiek na de commissie-Dijsselbloem*, Den Haag: 2014.
- Schaapman K., *Schoolstrijd: Ouders op de bres voor onderwijs*, Amsterdam: SWP 2000.
- Twist, M. van e.a., *Ibn Ghaldoun: Verhaal achter een bijzonder inspectiebezoek*, Den Haag: NSOB 2013.
- VROM Inspectie, *Compliance Assistance bij individuele toezichtbezoeken: de balans tussen controleren en ondersteunen*, Den Haag: 2006.

Kamerstukken

- Kamerstukken II 2005/06, 27 831, nr. 15. ('Minder last, meer effect').
- Kamerstukken II 2009/10, 32 193, nr. 2. (Wijziging WOT)
- Kamerstukken II 2009/10, 32193, nr. 3. (MvT)
- Kamerstukken II 2009/10, 32193, nr. 4. (Advies RvS)
- Kamerstukken II 2009/10, 32193, nr. 7. (Nota van wijziging)
- Kamerstukken II 2009/10, 32193, nr. 8. (Tweede nota van wijziging)
- Kamerstukken II 2009/10, 32193, nr. 14. (Antwoorden wijziging WOT)
- Handelingen II 2010/11, 94, 3. (Behandeling wijziging WOT)
- Handelingen II 2010/11, 103,6. (Voortgezette behandeling wijziging WOT)
- Handelingen I 2011/2012, 16, 9. (Behandeling EK)
- Handelingen I 2011/2012, 16, 11. (Voortgezette behandeling EK)

Website

<http://www.volkskrant.nl/dossier-de-beste-scholen/>